THE PRINCETON COLLOQUIUM

LECTURES ON MATHEMATICS

DELIVERED SEPTEMBER 15 TO 17, 1909, BEFORE MEMBERS OF THE AMERICAN MATHEMATICAL SOCIETY IN CONNECTION WITH THE SUMMER MEETING HELD AT PRINCETON UNIVERSITY, PRINCETON, N. J.

BY

GILBERT AMES BLISS

AND

EDWARD KASNER

NEW YORK
PUBLISHED BY THE
AMERICAN MATHEMATICAL SOCIETY
501 WEST 116TH STREET
1913

PRESS OF
THE NEW ERA PRINTING COMPANY
LANCASTER PA

PREFACE.

Soon after its expansion in 1894 into a national organization, the American Mathematical Society inaugurated the series of Colloquia which have been held in connection with its summer meetings since 1896, at intervals of two or three years. These Colloquia consist of courses of lectures delivered by specialists on selected chapters of their fields of work. Their purpose is to enable the members of the Society to keep in touch with the most recent advances of mathematical science and to stimulate a wide interest in its development.

The list of Colloquia thus far held is as follows:

I. THE BUFFALO COLLOQUIUM, 1896.

(*a*) Professor MAXIME BÔCHER, of Harvard University: "Linear Differential Equations, and Their Applications."

This Colloquium has not been published, but several papers appeared at about the time of the Colloquium, in which the author dealt with topics treated in the lectures.*

(*b*) Professor JAMES PIERPONT, of Yale University: "Galois's Theory of Equations."

Published in the *Annals of Mathematics*, series 2, volumes 1 and 2 (1900).

II. THE CAMBRIDGE COLLOQUIUM, 1898.

(*a*) Professor WILLIAM F. OSGOOD, of Harvard University: "Selected Topics in the Theory of Functions."

Published in the *Bulletin of the American Mathematical Society*, volume 5 (1898), pages 59–87.

* Two of these papers were: "Regular points of linear differential equations of the second order," Harvard University, 1896; "Notes on some points in the theory of linear differential equations," *Annals of Mathematics*, vol. 12 (1898).

(*b*) Professor ARTHUR G. WEBSTER, of Clark University: "The Partial Differential Equations of Wave Propagation."

III. THE ITHACA COLLOQUIUM, 1901.

(*a*) Professor OSKAR BOLZA, of the University of Chicago: "The Simplest Type of Problems in the Calculus of Variations."

Published in amplified form under the title: Lectures on the Calculus of Variations, Chicago, 1904.

(*b*) Professor ERNEST W. BROWN, of Haverford College: "Modern Methods of Treating Dynamical Problems, and in Particular the Problem of Three Bodies."

IV. THE BOSTON COLLOQUIUM, 1903.

(*a*) Professor HENRY S. WHITE, of Northwestern University: "Linear Systems of Curves on Algebraic Surfaces."

(*b*) Professor FREDERICK S. WOODS, of the Massachusetts Institute of Technology: "Forms of Non-Euclidean Space."

(*c*) Professor EDWARD B. VAN VLECK, of Wesleyan University: "Selected Topics in the Theory of Divergent Series and Continued Fractions."

This Colloquium was published for the Society in the volume: The Boston Colloquium Lectures on Mathematics, New York, Macmillan, 1905.

V. THE NEW HAVEN COLLOQUIUM, 1906.

(*a*) Professor ELIAKIM H. MOORE, of the University of Chicago: "On the Theory of Bilinear Functional Operations."

(*b*) Professor ERNEST J. WILCZYNSKI, of the University of California: "Projective Differential Geometry."

(*c*) Professor MAX MASON, of Yale University: "Selected Topics in the Theory of Boundary Value Problems of Differential Equations."

Published by Yale University in the volume: The New Haven Mathematical Colloquium, New Haven, Yale University Press, 1910.

VI. THE PRINCETON COLLOQUIUM, 1909.

(*a*) Professor GILBERT A. BLISS, of the University of Chicago: "Fundamental Existence Theorems."

(*b*) Professor EDWARD KASNER, of Columbia University: "Differential-Geometric Aspects of Dynamics."

This Colloquium is published here in full.

The Colloquia of the Society are to an extent comparable with the reports regularly presented to Section A of the British Association for the Advancement of Science and to the Deutsche Mathematiker-Vereinigung, and in so far play a rôle complementary to those of the *Bulletin* and *Transactions*. The Society will doubtless adopt the custom of publishing the lectures of each Colloquium in a corresponding volume.

The Seventh Colloquium will be held in connection with the twentieth summer meeting of the Society at Madison, Wisconsin during the week September 8–13, 1913. Courses of lectures will be given by Professor LEONARD E. DICKSON, of the University of Chicago, and Professor WILLIAM F. OSGOOD, of Harvard University. Thus for the first time an interval of four years has elapsed between successive Colloquia. As a suitable reflection and desirable stimulation of the mathematical activity of this country, it would seem desirable that the Colloquia should be held oftener. To avoid collision with the meetings of the International Congress of Mathematicians, the Colloquia might perhaps be arranged for every odd numbered year.

E. H. MOORE.

FUNDAMENTAL EXISTENCE THEOREMS

BY

GILBERT AMES BLISS

CONTENTS

CHAPTER III

Existence Theorems for Differential Equations

FUNDAMENTAL EXISTENCE THEOREMS

BY

GILBERT AMES BLISS

INTRODUCTION

The existence theorems to which these lectures are devoted have been the subject of a long sequence of investigations extending from the time of Cauchy to the present day, and have found application at the basis of a variety of mathematical theories including, as perhaps of especial importance, the theory of algebraic functions and the calculus of variations. If a single solution $(a; b) = (a_1, a_2, \cdots, a_m; b_1, b_2, \cdots, b_n)$ of a set of equations

$$f_\alpha(x_1, x_2, \cdots, x_m; y_1, y_2, \cdots, y_n) = 0 \quad (\alpha = 1, 2, \cdots, n)$$

is known, then in a neighborhood of $(a; b)$ there is one and only one other solution corresponding to each set of values x in a properly chosen neighborhood of the values a, and in the totality of solutions $(x; y)$ so defined the variables y are single-valued and continuous functions of the x's. If a set of initial constants $(\xi, \eta_1, \eta_2, \cdots, \eta_n)$ is given, then in a neighborhood of these values there is one and but one continuous arc

$$y_\alpha = y_\alpha(x) \qquad (\alpha = 1, 2, \cdots, n)$$

satisfying the differential equations

$$\frac{dy_\alpha}{dx} = g_\alpha(x, y_1, y_2, \cdots, y_n) \quad (\alpha = 1, 2, \cdots, n)$$

and passing through the initial values η when $x = \xi$.

2

The formulation and first satisfactory proofs of these theorems, at least for the case where only two variables x, y are involved, seem to be ascribed with unanimity to Cauchy. For the implicit functions his proof rested upon the assumption that the function f should be expressible by means of a power series, and the solution he sought was also so expressible, a restriction which was later removed with remarkable insight by Dini. For a differential equation, on the other hand, Cauchy assumed only the continuity of the function g and its first derivative for y, and his method of proof, with the well-known alteration due to Lipschitz, retains to-day recognized advantages over those of later writers.

In the following pages (§§ 1, 16) the two theorems stated above are proved with such alterations in the usual methods as seemed desirable or advantageous in the present connection. The proof given for the fundamental theorem of implicit functions is applicable when the independent variables x are replaced by a variable p which has a range of much more general type than a set of points in an m-dimensional x-space.* It is not necessary always to know an initial solution in order that others may be found. In the treatment of Kepler's equation, for example, which defines the eccentric anomaly of a planet moving in an elliptical orbit in terms of the observed mean anomaly, one starts with an approximate solution only and determines an exact solution by means of a convergent succession of approximations. This procedure is closely allied to a method of approximation due to Goursat (§ 3), suggested apparently by Picard's treatment of the existence theorem for differential equations.

One of the principal purposes of the paragraphs which follow, however, is to free the existence theorems as far as possible from

* The notion of a general range has been elucidated by Moore, The New Haven Mathematical Colloquium, page 4, the special cases which he particularly considers being enumerated on page 13. An application of the method of § 1 of these lectures when the range of p is a set of continuous curves, has been made by Fischer, "A generalization of Volterra's derivative of a function of a line," Dissertation, Chicago (1912).

the often inconvenient restriction which is implied by the words "in a neighborhood of," or which is so aptly expressed in German by the phrase "im Kleinen." It is evident from very simple examples that the totality of solutions $(x; y)$ associated continuously with a given initial solution of a system of equations $f = 0$ of the form described above, can not in general have the property that the variables y are everywhere single-valued functions of the variables x, and the result of attempting, perhaps unconsciously, to preserve the single-valued character of the solutions has been the restriction of the region to which the existence theorems apply. In order to avoid this difficulty and to characterize to some extent the totality of solutions associated continuously with a given initial one in a region specified in advance, the writer has introduced (§ 5) the notion of a particular kind of point set called a sheet of points. In a suitably chosen neighborhood of a point $(a; b)$ of the sheet there corresponds to every set of values x sufficiently near to the values a exactly one point $(x; y)$ of the sheet, and the single-valued functions y so determined are continuous and have continuous first derivatives. This condition does not at all imply that there are no other points of the sheet outside the specified neighborhood of the point $(a; b)$ and having a projection x near to a. With the help of the notion of a sheet of points it can be concluded that with any initial solution $(a; b)$ of the equations $f = 0$ there is associated a unique sheet S of solutions whose only boundary points are so-called exceptional points where the functions f either actually fail, or else are not assumed, to have the continuity and other properties which are demanded in the proof of the well-known theorem for the existence of solutions in a neighborhood of an initial one. It is important oftentimes to know whether or not a sheet of solutions is actually single-valued throughout its entire extent, and a criterion sufficient to ensure this property has also been derived (§ 7).

On the basis of these results some important theorems concerning the transformation of plane regions into regions of

another plane by means of equations of the form

$$x_1 = \psi_1(y_1, y_2), \qquad x_2 = \psi_2(y_1, y_2),$$

as in the theory of conformal transformation, have been deduced (§ 8). If the functions ψ have suitable continuity properties and a non-vanishing functional determinant in the interior of a simply closed regular curve B in the y-plane, and if B is transformed into a simply closed regular curve A of the x-plane, then the equations define a one-to-one correspondence between the interiors of A and B, and the inverse functions so defined have continuity properties similar to those of ψ_1 and ψ_2. This is but a sample of the theorems which may be stated. Others are also given (§ 8) which apply to the transformation of regions not necessarily finite, and to systems containing more than two equations.

The theory of the singularities of implicit functions is of considerable difficulty and has been but incompletely developed. For a transformation of the form above in which the functions ψ_1, ψ_2 are analytic, the singular point to be studied, at which the functional determinant $D = \partial(\psi_1, \psi_2)/\partial(y_1, y_2)$ vanishes, as well as its image in the x-plane, may both without loss of generality be supposed at the origin. The most general case under these circumstances is that for which the determinant D does not vanish identically and the equations $\psi_1 = 0$, $\psi_2 = 0$ have no real solutions in common near the origin except the values $y_1 = y_2 = 0$ themselves. It is found that the branches of the curve $D = 0$ bound off with a suitably chosen circle about the origin a number of triangular regions. Each of these regions is transformed in a one-to-one way into a sort of Riemann surface on the x-plane which winds about the origin and is bounded by the image of the boundary of the triangular region (see § 11, Fig. 6). If the signs of D in two adjacent triangular regions are opposite, then their images overlap along the common boundary; otherwise they adjoin without overlapping. At any point of one of the Riemann surfaces the inverse functions defined

by the transformation are continuous and in the interior of the surface they have everywhere continuous derivatives. These results are obtained by means of applications of the theorem described above for the transformation of the interior of a simply closed curve B; and the same method of procedure would undoubtedly be of service when the curves $\psi_1 = 0$, $\psi_2 = 0$ have real branches through the origin in common, which must occur whenever they have common points in every neighborhood of the values $y_1 = y_2 = 0$. The case where the determinant D vanishes identically is also considered (§ 12).

For the singularities of implicit functions defined by a system of equations $f = 0$ there is a generalization of the preparation theorem of Weierstrass (§ 9) suggested to the writer by some remarks in the introduction of Poincaré's Thesis, and by a study of the elimination theory of Kronecker for algebraic equations. The theorem is presented here (§ 13) for two equations and two variables y_1, y_2 in the form originally given at the time of the Princeton Colloquium, but the method of proof is similar to that of a later paper* and applies with suitable modifications to a system containing more equations and independent variables. These results can not by any means be said to afford a complete characterization of the singularities of implicit functions, but it is hoped that they may be useful in paving the way for researches of a more comprehensive character.

The writer published some years ago a paper† concerning the extensibility of the solutions of a system of differential equations, of the form specified above, from boundary to boundary of a finite closed region R in which the functions g_a are supposed to have suitable continuity properties. In the last chapter of these lectures the character of the region has been generalized so that no restrictions as to its finiteness or closure are made, and it is shown that the approximations of Cauchy converge to a solution over an interval

* See the footnote to page 73.

† "The solutions of differential equations of the first order as functions of their initial values," *Annals of Mathematics*, 2d series, vol. 6 (1904), page 49.

in the interior of which the limiting curve is continuous and interior to R, while at the ends of the interval the only limit points of the curve are at infinity or else are on the boundary of the region. The solutions so defined are continuous and differentiable with respect to their initial values, a property which once proved is of great service in many of the applications of the existence theorems. One situation in which these results have an important bearing is related to a partial differential equation of the first order

$$F(x, y, z, \partial z/\partial x, \partial z/\partial y) = 0.$$

When this equation is analytic, any analytic curve C, which is not a so-called integral curve, defines uniquely an analytic surface containing the curve and satisfying the differential equation. The uniqueness in this case is a consequence, in the first place, of the fact that an analytic surface is completely determined when an initial series defining its values in a limited region is given, and, in the second place, of the theorem that at a given point and normal of the initial curve C satisfying the differential equation there is but one series defining an integral surface including the points of C and having the given initial normal. It is not self evident in what sense a solution of a non-analytic equation is uniquely determined by an initial curve, as may be seen by very simple examples. An initial curve which is not an integral curve will in general have associated with it, however, a strip of normals which satisfy the partial differential equation, and whose elements as initial values determine a one-parameter family of characteristic strips simply covering a region R_{xy} of the xy-plane about the projection of the initial curve C. There is one and but one integral surface of the differential equation with a continuously turning tangent plane and continuous curvature, which is defined at every point of the region R_{xy} and contains the initial curve C and its strip of normals (§ 19).

CHAPTER I

ORDINARY POINTS OF IMPLICIT FUNCTIONS

§ 1. The Fundamental Theorem

The fundamental theorem of the implicit function theory states the existence of a set of functions

$$y_\alpha = y_\alpha(x_1, x_2, \cdots, x_m) \quad (\alpha = 1, 2, \cdots, n)$$

which satisfy a system of equations of the form

$$(1) \qquad f_\alpha(x_1, x_2, \cdots, x_m; y_1, y_2, \cdots, y_n) = 0 \quad (\alpha = 1, 2, \cdots, n)$$

in a neighborhood of a given initial solution $(a; b)$. Dini's method,* for the case in which the functions f are only assumed to be continuous and to have continuous first derivatives, is to show the existence of a solution of a single equation, and then to extend his result by mathematical induction to a system of the form given above, a plan which has been followed, with only slight alterations and improvements in form, by most writers on the theory of functions of a real variable. In a more recent paper† Goursat has applied a method of successive approximations which enabled him to do away with the assumption of the existence of the derivatives of the functions f with respect to the independent variables x.

One can hardly be dissatisfied with either of these methods of attack. It is true that when the theorem is stated as precisely as in the following paragraphs, the determination of the neighborhoods at the stage when the induction must be made is rather inelegant, but the difficulties encountered are not serious. The introduction of successive approximations is an interesting step,

* Lezioni di Analisi infinitesimale, vol. 1, chap. 13. For historical remarks, see Osgood, Encyclopädie der mathematischen Wissenschaften, II, B 1, § 44 and footnote 30.

† *Bulletin de la Société mathématique de France*, vol. 31 (1903), page 185.

though it does not simplify the situation and indeed does not add generality with regard to the assumptions on the functions f. The method of Dini can in fact, by only a slight modification, be made to apply to cases where the functions do not have derivatives with respect to the variables x. The proof which is given in the following paragraphs seems to have advantages in the matter of simplicity over either of the others. It applies equally well, without induction, to one or a system of equations, and requires only the initial assumptions which Goursat mentions in his paper.

Where it is possible without sacrificing clearness, the row letters f, x, y, p, a, b will be used to denote the systems

$$\begin{aligned} f &= (f_1, f_2, \cdots, f_n), & x &= (x_1, x_2, \cdots, x_m), \\ y &= (y_1, y_2, \cdots, y_n), & a &= (a_1, a_2, \cdots, a_m), \\ b &= (b_1, b_2, \cdots, b_n), & p &= (a_1, a_2, \cdots, a_m; b_1, b_2, \cdots, b_n). \end{aligned}$$

In this notation the equations (1) have the form

$$f(x; y) = 0,$$

the interpretation being that every element of f is a function of $x_1, x_2, \cdots, x_m; y_1, y_2, \cdots, y_n$, and every f_i is to be set equal to zero. The notations p_ϵ, a_ϵ, b_ϵ represent respectively the neighborhoods

$$|x - a| < \epsilon, \quad |y - b| < \epsilon; \quad |x - a| < \epsilon; \quad |y - b| < \epsilon$$

of the points p, a, b.

With these notations in mind the fundamental theorem which is to be proved may be stated as follows:

Hypotheses:

1) *the functions $f(x; y)$ are continuous, and have first partial derivatives with respect to the variables y which are also continuous, in a neighborhood of the point $(a; b)$ which will be denoted by p;*

2) $f(a; b) = 0$;

3) *the functional determinant $D = \partial(f_1, f_2, \cdots, f_n)/\partial(y_1, y_2, \cdots, y_n)$ is different from zero at p.*

Conclusions:

1) *a neighborhood p_ϵ can be found in which there corresponds to a given value x at most one solution $(x; y)$ of the equations $f(x; y) = 0$;*

2) *for any neighborhood p_ϵ with the property just described a constant $\delta \leqq \epsilon$ can be found such that every x in a_δ has associated with it a point $(x; y)$ which satisfies the equations $f(x; y) = 0$;*

3) *the functions $y(x_1, x_2, \cdots, x_m)$ so found are continuous in the region a_δ.*

For the neighborhood p_ϵ let one be chosen in which the continuity properties of the functions f are preserved. If $(x; y)$ and $(x; y')$ are two points in p_ϵ, it follows, by applying Taylor's formula to the differences $f(x; y') - f(x; y)$, that

$$f_1(x; y') - f_1(x; y) = \frac{\partial f_1}{\partial y_1}(y_1' - y_1) + \cdots + \frac{\partial f_1}{\partial y_n}(y_n' - y_n),$$

$$\cdots\cdots\cdots\cdots$$

$$f_n(x; y') - f_n(x; y) = \frac{\partial f_n}{\partial y_1}(y_1' - y_1) + \cdots + \frac{\partial f_n}{\partial y_n}(y_n' - y_n),$$

where the arguments of the derivatives $\partial f_\alpha / \partial y_\beta$ have the form $x; y + \theta_\alpha(y' - y)$, and $0 < \theta_\alpha < 1$. The determinant of these derivatives is different from zero when $(x; y') = (x; y) = (a; b)$, and hence must remain different from zero if p_ϵ is restricted so that in it the functional determinant D remains different from zero. It is then impossible that $(x; y)$ and $(x; y')$ should both be solutions of the equations $f(x; y) = 0$, if y is distinct from y'.

In the corresponding region b_ϵ the function

$$\varphi(a; y) = f_1^2(a; y) + f_2^2(a; y) + \cdots + f_n^2(a; y)$$

has a minimum for $y = b$, since for that value it vanishes and for every other it is positive. In particular

$$\varphi(a; \eta) - \varphi(a; b) > m > 0$$

when η ranges over the *closed* set of points η forming the boundary of b_ϵ, on account of the continuity of φ, and the inequality

$$\varphi(x; \eta) - \varphi(x; b) > m$$

remains true for all values x in a suitably chosen domain a_δ. Hence for a fixed x in a_δ the minimum of $\varphi(x; y)$ is attained at a point y *interior* to b_e. At such a point, however,

$$\frac{1}{2}\frac{\partial\varphi}{\partial y_1} = f_1\frac{\partial f_1}{\partial y_1} + f_2\frac{\partial f_2}{\partial y_1} + \cdots + f_n\frac{\partial f_n}{\partial y_1} = 0,$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\frac{1}{2}\frac{\partial\varphi}{\partial y_n} = f_1\frac{\partial f_1}{\partial y_n} + f_2\frac{\partial f_2}{\partial y_n} + \cdots + f_n\frac{\partial f_n}{\partial y_n} = 0,$$

and this can happen only when all the elements of f are zero, since the functional determinant D is different from zero in p_e. It follows that to every point x in a_δ there corresponds in p_e a solution $(x; y)$ of the equations $f(x; y) = 0$.

The functions $y(x_1, x_2, \cdots, x_m)$ defined in this way over the region a_δ are all continuous. For consider the values y and $y + \Delta y$ corresponding to two points x and $x + \Delta x$. By applying Taylor's formula it follows from the relations

$$f(x; y + \Delta y) - f(x; y) = f(x; y + \Delta y) - f(x + \Delta x; y + \Delta y),$$

which are true because $(x; y)$ and $(x + \Delta x; y + \Delta y)$ both make $f = 0$, that

$$\frac{\partial f_1}{\partial y_1}\Delta y_1 + \frac{\partial f_1}{\partial y_2}\Delta y_2 + \cdots + \frac{\partial f_1}{\partial y_n}\Delta y_n$$
$$= f_1(x; y + \Delta y) - f_1(x + \Delta x; y + \Delta y),$$

(2) $$\cdots\cdots\cdots\cdots\cdots$$

$$\frac{\partial f_n}{\partial y_1}\Delta y_1 + \frac{\partial f_n}{\partial y_2}\Delta y_2 + \cdots + \frac{\partial f_n}{\partial y_n}\Delta y_n$$
$$= f_n(x; y + \Delta y) - f_n(x + \Delta x; y + \Delta y),$$

where the arguments of the derivatives $\partial f_\alpha/\partial y_\beta$ have the form $x; y + \theta_\alpha\Delta y$ $(0 < \theta_\alpha < 1)$. The determinant of these derivatives is different from zero on account of the way in which p_e was chosen, and the second members of the equations approach zero with Δx. Hence the same must be true of the quantities

Δy, and thus the functions $y(x_1, x_2, \cdots, x_m)$ are seen to be continuous.

A similar application of Taylor's formula leads to the conclusion:

If the functions f have derivatives of the first order with respect to x_k which are continuous in the neighborhood of p, so have also the functions $y(x_1, x_2, \cdots, x_m)$ in the region a_δ; and if the f's have all derivatives of the nth order continuous, so have the functions $y(x_1, x_2, \cdots, x_m)$.

For suppose

$$\Delta x_1 \neq 0, \quad \Delta x_2 = \Delta x_3 = \cdots = \Delta x_m = 0.$$

Then by applying Taylor's formula to the second members of equations (2) it follows that

$$\begin{aligned}
&\frac{\partial f_1}{\partial y_1}\frac{\Delta y_1}{\Delta x_1} + \frac{\partial f_1}{\partial y_2}\frac{\Delta y_2}{\Delta x_1} + \cdots + \frac{\partial f_1}{\partial y_n}\frac{\Delta y_n}{\Delta x_1} + \frac{\partial f_1'}{\partial x_1} = 0,\\
&\quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad .\\
&\frac{\partial f_n}{\partial y_1}\frac{\Delta y_1}{\Delta x_1} + \frac{\partial f_n}{\partial y_2}\frac{\Delta y_2}{\Delta x_1} + \cdots + \frac{\partial f_n}{\partial y_n}\frac{\Delta y_n}{\Delta x_1} + \frac{\partial f_n}{\partial x_1} = 0,
\end{aligned}$$

where the arguments of the derivatives $\partial f_a/\partial x_1$ have the form $x + \theta_a'\Delta x;\ y + \Delta y$. Hence as Δx_1 approaches zero the quotients $\Delta y_a/\Delta x_1$ approach limits $\partial y_a/\partial x_1$ which satisfy the equations

$$(3)\qquad \begin{aligned}
&\frac{\partial f_1}{\partial y_1}\frac{\partial y_1}{\partial x_1} + \frac{\partial f_1}{\partial y_2}\frac{\partial y_2}{\partial x_1} + \cdots + \frac{\partial f_1}{\partial y_n}\frac{\partial y_n}{\partial x_1} + \frac{\partial f_1}{\partial x_1} = 0,\\
&\quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad .\\
&\frac{\partial f_n}{\partial y_1}\frac{\partial y_1}{\partial x_1} + \frac{\partial f_n}{\partial y_2}\frac{\partial y_2}{\partial x_1} + \cdots + \frac{\partial f_n}{\partial y_n}\frac{\partial y_n}{\partial x_1} + \frac{\partial f_n}{\partial x_1} = 0,
\end{aligned}$$

where the arguments of the derivatives of f are now $(x; y)$. A similar consideration shows the existence of the first derivatives with respect to the variables $x_2, x_3, \cdots, x_m$. The existence of the higher derivatives follows from the observation that the solutions of equations (3) for the quotients $\partial f_a/\partial y_\beta$ are

differentiable $n-1$ times with respect to the variables x, on account of the assumption that the functions f are differentiable n times.

§ 2. Equations in which the Functions are Analytic

It seems necessary to proceed differently in order to prove that when the functions f in equations (1) are analytic with coefficients and variables permitted to assume imaginary values, the solutions $y = y(x_1, x_2, \cdots, x_m)$ are also analytic functions of the variables x. The following theorem can first be proved:

When the functions f are formal series in the variables x; y with literal coefficients and having no constant terms, then there exists one and but one set of series

$$y_\alpha = y_\alpha(x_1, x_2, \cdots, x_m) \tag{4}$$

for the variables y, which vanish with the x's and satisfy identically the equations $f(x; y) = 0$. Each coefficient in the series y is rational in a finite number of those of the functions f, the only denominators occurring being powers of the determinant R of the coefficients of the linear terms in y.

To prove this let the equations $f = 0$ be written in the form

$$\begin{gathered} a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n = g_1(x; y), \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nn}y_n = g_n(x; y), \end{gathered}$$

where the functions g have no linear terms in y. By multiplying these equations by proper factors and adding, they may be made to take the form

$$y_\alpha = h_\alpha(x; y) \qquad (\alpha = 1, 2, \cdots, n), \tag{5}$$

where the series h have still no linear terms in y and have coefficients which are rational in those of the functions f, the only denominators occurring being the determinant R. Any series for y which satisfy formally the original equations must satisfy the last equations, and vice versa.

Consider now a set of series (4) in which the coefficients are indeterminates c. If they satisfy the equations (5) identically, then by comparison of coefficients on the two sides it is seen that any coefficient c_ν of a term of degree ν must be equal to a polynomial, with positive integral coefficients, in a finite number of the coefficients of the functions h and in the coefficients $c_{\nu-k}$ of terms in the functions y of lower degree than ν. For there are at most a finite number of terms on the right of any given degree ν, and since the functions h have no linear terms in the variables y it follows that wherever the term containing c_ν occurs it is always multiplied by a y or by a power of some of the variables x, and hence c_ν can only appear in terms of degree greater than ν. Since the coefficients of the linear terms in the functions y are equal respectively to corresponding coefficients in the functions h, it follows by an easy induction that every coefficient in the functions y must be a polynomial with positive integral coefficients in a finite number of the coefficients of the functions h. There is evidently but one set of series (4) of the kind described satisfying formally the equations (5), or what is the same thing, the equations $f = 0$.

For any numerical choice of the coefficients of the functions f in the domain of real or imaginary numbers for which the series f converge and the determinant $R = | a_{\alpha\beta} |$ is different from zero, the series (4) *for y will also be well-determined and convergent.*

For, a set of equations

$$y_\alpha = H_\alpha(x; y) \qquad (\alpha = 1, 2, \cdots, n) \tag{6}$$

can be constructed whose coefficients are all positive and greater numerically than the corresponding coefficients in the functions h, and for which the corresponding series $y = Y(x_1, x_2, \cdots, x_m)$ converge. The coefficients in the functions Y will be greater numerically than the corresponding coefficients of the series $y(x_1, x_2, \cdots, x_m)$, and hence the series y will also converge.

To show this suppose that ρ is a positive constant smaller than the radii of convergence of the functions $h(x; y)$. Then

the series $h(\rho;\rho)$ are convergent, and each term is numerically smaller than a constant M chosen greater than the sum of the absolute values of the terms in any one of the series $h(\rho;\rho)$. The coefficient of any term in $h(x;y)$ is less than M/ρ^ν where ν is the degree of the term. The series

$$H_\alpha(x;y) = \frac{M}{\left(1-\frac{x_1+x_2+\cdots+x_m}{\rho}\right)\left(1-\frac{y_1+y_2+\cdots+y_n}{\rho}\right)} - M - M\frac{y_1+y_2+\cdots+y_n}{\rho}$$

are similar to the series $h(x;y)$ in the matter of missing terms, and dominate them in the manner described above, since the coefficient of any term of degree ν is M/ρ^ν or greater.

The unique series satisfying equations (6) will evidently be convergent if a convergent series u in x can be determined satisfying

$$u = \frac{M}{\left(1-\frac{x_1+x_2+\cdots+x_m}{\rho}\right)\left(1-\frac{nu}{\rho}\right)} - M - M\frac{nu}{\rho},$$

for then every series y can be put equal to that series u. The latter equation is however a quadratic in u and has the solution

$$u = \frac{\rho^2}{2n(\rho+Mn)}\left\{1-\sqrt{1-\frac{4Mn(\rho+Mn)}{\rho^2}\,\frac{\frac{x_1+x_2+\cdots+x_m}{\rho}}{1-\frac{x_1+x_2+\cdots+x_m}{\rho}}}\right\}$$

vanishing with x. This will certainly be representable by a convergent series in x provided that

$$|x_i| < \frac{\rho^3}{m(\rho+2Mn)^2} \qquad (i = 1, 2, \cdots, m),$$

since then the second term under the radical is numerically less than unity.

The two theorems which have just been proved enable one to make the following statement concerning the solutions whose existence was proved in § 1:

If the functions $f(x; y)$ are analytic in the region p_ϵ, then the solutions (4) of the equations $f(x; y) = 0$ are analytic at every point of the region a_δ.

It is only necessary to transform the origin of coordinates to the particular point $(x; y)$ of the solution which it is desired to investigate.

Furthermore when the domain in which the equations $f = 0$ are to be studied is the domain of complex numbers, a theorem analogous to that of § 1 may be stated.

If in the domain of complex numbers the functions $f(x; y)$ are analytic at a point $p(a; b)$ at which

$$f(a; b) = 0, \quad D(a; b) = \left[\frac{\partial(f_1, f_2, \cdots, f_n)}{\partial(y_1, y_2, \cdots, y_n)}\right]_{\substack{x=a \\ y=b}} \neq 0,$$

then there exists a neighborhood p_ϵ in which any x corresponds to at most one solution $(x; y)$, either real or complex, of the equations $f(x; y) = 0$. For any such choice of p_ϵ a neighborhood a_δ ($\delta \leq \epsilon$) can be found such that every point x in a_δ has associated with it a solution $(x; y)$ of the equations $f = 0$ in p_ϵ, and the values y for these solutions are defined by a set of functions

$$(7) \qquad y_\alpha = y_\alpha(x_1, x_2, \cdots, x_m) \qquad (\alpha = 1, 2, \cdots, n)$$

which are expressible as series in the differences $x - a$ convergent in the region a_δ.

The existence of the neighborhood p_ϵ is provable by the argument used in § 1, since for any two points $(x; y)$ and $(x; y')$ in the common domain of convergence of the functions f, equations of the form

$$f_\alpha(x; y') - f_\alpha(x; y) = A_{\alpha 1}(y_1' - y_1) + \cdots + A_{\alpha n}(y_n' - y_n), \qquad (\alpha = 1, 2, \cdots, n)$$

hold, where the coefficient $A_{\alpha\beta}$ is a convergent series in the dif-

ferences $x - a$, $y - b$, $y' - b$ with constant term equal to $a_{\alpha\beta}$. The existence of the coefficients A can be established by considering two analogous terms in $f(x; y)$ and $f(x; y')$. The difference of such a pair of terms will always be linearly expressible in terms of the differences

$$(y_\alpha' - b_\alpha) - (y_\alpha - b_\alpha) = y_\alpha' - y_\alpha \quad (\alpha = 1, 2, \cdots, n).$$

Furthermore for $(x, y, y') = (a, b, b)$ the derivative of the first member with respect to y_β' reduces to $a_{\alpha\beta}$, while that of the second is the constant term in $A_{\alpha\beta}$. Hence for these values of the variables the determinant $|A_{\alpha\beta}|$ reduces to $D(a, b) \neq 0$.

By transforming the origin of coordinates to the point (a, b) and applying the first two theorems of this section, it follows that there exists a set of convergent series (7) satisfying the equations $f = 0$ identically; and for a sufficiently small region a_δ the points $(x; y)$ which they define will all lie in the neighborhood p_ϵ.

§ 3. Goursat's Method of Approximation

The method of approximation which is to be presented in the following paragraphs is of interest primarily because it affords a direct method of finding the values of implicit functions, and justifies computations sometimes used in the applications of the theory. In order to exhibit this method suppose again that the functions f have the properties described in the principal theorem of § 1, and consider the following set of equations suggested by Taylor's formula:

$$(8)\quad \begin{aligned} f_1(x\,;y) + a_{11}(y_1' - y_1) + a_{12}(y_2' - y_2) + \cdots \\ + a_{1n}(y_n' - y_n) = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f_n(x\,;y) + a_{n1}(y_1' - y_1) + a_{n2}(y_2' - y_2) + \cdots \\ + a_{nn}(y_n' - y_n) = 0, \end{aligned}$$

in which the coefficient $a_{\alpha\beta}$ is the value of $\partial f_\alpha / \partial y_\beta$ at the point p. When solved for the variables y', these equations take the form

$$(9)\qquad y_\alpha' = \varphi_\alpha(x\,;y) \qquad (\alpha = 1, 2, \cdots, n),$$

and one verifies readily by substitution of these expressions in equations (8) that the functions φ and all of their first derivatives with respect to the elements of y are continuous near p; and at the point p itself φ_α has the value b_α, while all of its derivatives with respect to the y's vanish.

A sequence of systems $y^{(k)} = (y_1^{(k)}, y_2^{(k)}, \cdots, y_n^{(k)})$ beginning with the set

$$y' = [\varphi_1(x; b),\ \varphi_2(x; b),\ \cdots,\ \varphi_n(x; b)]$$

can now be defined by means of the recursion formulas (9), which are equivalent to

$$y_\alpha^{(k)} = \varphi_\alpha(x; y^{(k-1)}) \qquad (\alpha = 1, 2, \cdots, n).$$

Let p_ϵ be any neighborhood of p in which the continuity properties of f are retained, and in which the derivatives of φ remain numerically less than θ/n where $0 < \theta < 1$. If the values of x are restricted to a region $a_\delta(\delta \leqq \epsilon)$ so small that every element of the set y' satisfies the inequality

$$(10) \qquad | y_\alpha' - b_\alpha | < \epsilon(1 - \theta),$$

then the points $(x; y^{(k)})$ will all lie in the neighborhood p_ϵ and will approach uniformly a limiting point $(x; y)$ which is a solution of the equations (1).

To prove these statements one needs only to apply successively the inequality

$$\begin{aligned} | y_\alpha^{(k)} - y_\alpha^{(k-1)} | &= | \varphi_\alpha(x\,; y^{(k-1)}) - \varphi_\alpha(x\,; y^{(k-2)}) | \\ &\leqq \frac{\theta}{n}\{ | y_1^{(k-1)} - y_1^{(k-2)} | + | y_2^{(k-1)} - y_2^{(k-2)} | \\ &\qquad + \cdots + | y_n^{(k-1)} - y_n^{(k-2)} | \}, \end{aligned}$$

which follows readily by an application of Taylor's formula. Since the inequalities (10) hold, the last formula successively applied shows that

$$| y_\alpha^{(k)} - y_\alpha^{(k-1)} | \leqq \theta^{k-1}\epsilon(1 - \theta).$$

Consequently the sum $y_\alpha^{(k)}$ of the first $k + 1$ terms of the series

$$(11) \quad b_\alpha + (y_\alpha' - b_\alpha) + (y_\alpha'' - y_\alpha') + \cdots + (y_\alpha^{(k)} - y_\alpha^{(k-1)}) + \cdots$$

3

differs in absolute value from b_α by a quantity which is less than

$$\epsilon(1-\theta)(1+\theta+\theta^2+\cdots+\theta^{k-1}) = \epsilon(1-\theta^k) < \epsilon.$$

Hence the points $(x; y)$ all lie in the neighborhood p_ϵ, and the series (11) is uniformly convergent in the neighborhood a_δ.

The limiting point $(x; y)$ evidently satisfies the equations $f = 0$. For at every stage the values (x, y, y') $(x, y^{(k-1)}, y^{(k)})$ satisfy the equations (8), and the first members of these equations approach uniformly the values $f(x; y)$.

The process of determining the solutions described above is evidently one of trial and error. The values $y = b$ being first substituted, the equations (9) determine approximately the correction $y' - b$ which must be added to b in order to obtain a solution for any value of x near to a. For the values so corrected the equations (9) give again a new correction $y'' - y'$, and so on.

It is ordinarily presupposed that an initial solution $(a; b)$ is given, *but the process may also lead to the discovery of a solution in case only an initial point which approximately satisfies the equation is known.* To show this suppose that the functions f are continuous and have continuous first partial derivatives with respect to the variables y in a closed region R of points $(x; y)$ in which the functional determinant $D(x; y)$ is different from zero. The functions φ in equations (9) are to be thought of as depending upon (x, y), and also upon the variables $(a; b)$ which enter in the derivatives $a_{\alpha\beta}$. Then the expressions $\varphi(x, y, a, b)$, $\varphi_y(x, y, a, b)$ are continuous when $(x; y)$, $(a; b)$ lie in R, and all of the derivatives φ_y vanish identically when $(x; y) = (a; b)$. The value of $\varphi(a, b, a, b)$ is not necessarily b, however, when $(a; b)$ is not a solution. Two positive constants, $\theta < 1$ and ϵ, can be determined so that

$$|\varphi_y(x, y, a, b)| < \theta/n$$

whenever $(a; b)$ and $(x; y)$ satisfy the inequalities

$$|x-a| < \epsilon, \qquad |y-b| < \epsilon.$$

If now there exists a point $p(a; b)$ for which the neighborhood p_ϵ

is entirely within R, and such that

$$| \varphi(a, b, a, b) - b | < \epsilon(1 - \theta),$$

then the sequence $y^{(k)}$ defined converges uniformly as before in a neighborhood a_δ of the point a and determines a solution $(x; y)$.

As an example consider the equation

$$y - e \sin y = x \qquad (0 < e < 1), \tag{12}$$

which in the theory of elliptic orbits determines the value of the eccentric anomaly y in terms of the mean anomaly x. The function φ is in this case

$$\varphi(x, y, a, b) = \frac{e(\sin y - y \cos b) + x}{1 - e \cos b}$$

and φ_y remains less than θ when

$$| y - b | < \theta \frac{1 - e}{e} = \epsilon.$$

For any given $x = a$, a value $y = b$ can be determined, by graphical methods for example, so that

$$| \varphi(a, b, a, b) - b | = \left| \frac{b - e \sin b - a}{1 - e \cos b} \right| < \theta \frac{1 - e}{e} (1 - \theta).$$

The process described above therefore converges in a suitably chosen neighborhood of $x = a$, and a solution of equation (12) can be found when an approximate solution only has been determined in advance.

§ 4. Bolza's Extension of the Fundamental Theorem*

The neighborhood P_ϵ of a set of points P in the space $(x; y)$ is the totality of points $(x; y)$ which satisfy inequalities of the form

$$| x - a | < \epsilon, \qquad | y - b | < \epsilon,$$

* Vorlesungen über Variationsrechnung, page 160: also *Mathematische Annalen*, vol. 63 (1906), page 247. The theorem was proved independently by Mason and Bliss, "Fields of extremals in space," *Transactions of the American Mathematical Society*, vol. 11 (1910), page 326.

where $(a; b)$ is some point of P. The sets of points (a) and (b) which belong to points $(a; b)$ of P are the projections of P in the x- and y-spaces, and will be denoted by A and B, respectively.

The fundamental theorem of § 1 *remains true if in its statement the single point p is replaced by a set of points P which is finite and closed, and which furthermore has the property that no two distinct points $(a; b)$, $(a'; b')$ of P have the same projection $a' = a$. According to the conclusions of the theorem there exists then a neighborhood P_ϵ in which no two solutions of the equations $f(x; y) = 0$ have the same projection x, and a neighborhood A_δ in which every x surely belongs to a solution $(x; y)$ in P_ϵ. The single-valued functions $y(x_1, x_2, \cdots, x_m)$ so defined in A_δ are continuous, and if the functions $f(x; y)$ have continuous derivatives of the n-th order in a neighborhood of P, so have the functions $y(x_1, x_2, \cdots, x_m)$ in A_δ.*

To prove the theorem suppose first that a sequence of positive constants ϵ_k $(k = 1, 2, \cdots)$ approaching zero has been selected arbitrarily. If the first part of the theorem were not true, then in any neighborhood P_{ϵ_k} there would be two distinct solutions $(x; y)_k$ and $(x; y')_k$ of the equations $f(x; y) = 0$, which would satisfy, respectively, inequalities of the form

$$(13) \qquad \begin{aligned} &| x - \alpha | < \epsilon_k, \quad | y - \beta | < \epsilon_k; \\ &| x - \alpha' | < \epsilon_k, \quad | y' - \beta' | < \epsilon_k \end{aligned}$$

with two points $(\alpha; \beta)_k$ and $(\alpha'; \beta')_k$ of the set P. Since P is finite and closed, the sequence of values $(\alpha, \beta; \alpha', \beta')_k$ has a point of condensation $(a, b; a', b')$ for which $(a; b)$ and (a', b') are both in P. From the inequalities (13) it follows that $(a, b; a', b')$ is also a point of condensation for the sequence $(x, y; x, y')_k$, and therefore a and a' must be the same. The values b and b' must also be identical since P contains only one point $p(a; b)$ with the projection a. According to the original statement of the fundamental theorem in § 1, a neighborhood p_ϵ can be chosen in which no two solutions of the equations $f(x; y) = 0$ have the same projection x. Hence the existence of the sequences $(x; y)_k$ and $(x; y')_k$ with the common point of

condensation $(a; b)$ is contradicted, and it must always be possible to select a neighborhood P_ϵ in which distinct solutions of the equations $f = 0$ always have distinct projections x.

A similar argument shows that a neighborhood A_δ can be selected so that to any point of it there corresponds a solution of the equations $f = 0$. Otherwise to each δ_k of a sequence of constants approaching zero, there would correspond a point $(x)_k$ in the region A_{δ_k} which would belong to no solution in P_ϵ. To each $(x)_k$ there would correspond a point $(\alpha)_k$ in A satisfying the inequalities

$$| x - a | < \delta_k$$

with the values $(x)_k$, and the points $(\alpha)_k$ would have a point of condensation a in A, which would also be a point of condensation for the sequence $(x)_k$, since A is finite and closed when P is so. But by the original theorem of § 1, again, it is known that a neighborhood a_δ of a can be chosen in which every point x has associated with it a solution $(x; y)$ in p_ϵ, where $p(a; b)$ is the point of P having the projection a. Consequently the existence of the sequence $(x)_k$ is contradicted.

If now the region P_ϵ is so restricted that the functional determinant $D(x; y)$ remains different from zero throughout it, then the original theorem of § 1 can be applied to show that the functions $y(x_1, x_2, \cdots, x_n)$ are continuous at any point of the region A_δ and possess as many continuous derivatives as are possessed by the functions $f(x; y)$.

§ 5. The Unique Sheet of Solutions Associated with an Initial Solution

The points of the space $(x; y)$ may be divided into two classes, ordinary points and exceptional points, with respect to the functions f. An *ordinary point* is one at which the first and third hypotheses of the theorem of § 1 are postulated, that is, one near which the functions f and their first derivatives with respect to y are continuous and the functional determinant $D = \partial(f_1, f_2, \cdots,$

$f_n)/\partial(y_1, y_2, \cdots, y_n)$ is different from zero. An *exceptional point* is one at which some of these conditions are not fulfilled or are not presupposed.

A *sheet of points* in the $(m+n)$-dimensional space $(x; y)$ may be defined as a point set S with the property that for any point $p(a; b)$ belonging to the set a neighborhood p_ϵ can always be found such that no two points of S in p_ϵ have the same projection x. In other words, the variables y are single-valued functions $y(x_1, x_2, \cdots, x_m)$ in the neighborhood of the point p, for points of the sheet.

If for any neighborhood b_ϵ of the kind just described, a region a_δ $(\delta \leqq \epsilon)$ can be found in which every point x belongs to a point of S in p_ϵ, then p is said to be an *interior point* of the sheet S.

A *boundary point* is a limit point of points of the sheet, which is not itself an interior point and may not even belong to S.

A sheet is said to be *connected* if every pair $(x'; y')$, $(x''; y'')$ of its interior points can be joined by a continuous curve

$$x = x(t), \qquad y = y(t) \qquad (t' \leqq t \leqq t''),$$

consisting entirely of interior points of the sheet.

In the following pages it is always to be understood that the sheets considered are continuous and have continuous first derivatives, or in other words at any interior point of one of them the functions $y(x_1, x_2, \cdots, x_m)$ mentioned above have these properties. A sheet will be said to become infinite near a point x' if x' is the limit of the projections of a sequence of points $(x; y)$ of the sheet for which one at least of the variables y approaches infinity.

With the preceding agreements as to nomenclature in mind, it is possible to prove the following theorem:

If a point $p(a; b)$ is an ordinary point for the functions f and satisfies the equations $f = 0$, then there passes through p one and only one connected sheet of solutions of these equations, with the properties:

1) *all points of the sheet are ordinary points of the functions f;*
2) *all points are interior points;*

3) *the only boundary points of the sheet are exceptional points for the system f.*

The set of points

$$[x_1, x_2, \cdots, x_m;\ y_1(x_1, x_2, \cdots, x_m), \quad \cdots, \quad y_n(x_1, x_2, \cdots, x_m)]$$

defined over the region a_δ by the principal theorem of § 1, is a sheet S_1 of solutions of the equations $f = 0$ which satisfies all the requirements of the theorem just stated except possibly the last. Its points are all interior points since the region a_δ is defined by inequalities only. If any boundary point $p'(a'; b')$ of S_1 is an ordinary point of the functions f it must satisfy the equations $f = 0$, since the f's are continuous and p' is a limit point of points on S_1. Consequently the theorem of § 1 can be applied in the neighborhood of p', and the sheet S' so determined near p' forms with S_1 a new set S_2. This process may be repeated any number of times, and the totality of points which can be attained by a finite number of such extensions, constitutes the sheet S required in the theorem.

The set of points S so determined constitutes a sheet, since any point q of it is an ordinary point and a solution of the equations $f = 0$, and according to the theorem of § 1 the solutions of these equations in the neighborhood of q have the property which is characteristic of a sheet. From the manner of its construction the sheet is evidently connected and consists entirely of interior points. If any boundary point q of S were an ordinary point of the functions f, the sheet could be extended to include q as an interior point by the process described in the preceding paragraph.

There could not be a second sheet Σ containing a point π not in S and having the properties stated in the theorem. For there would in that case be a continuous curve

$$x = x(t), \quad y = y(t) \qquad (t_1 \leqq t \leqq t_2)$$

in Σ joining p with π and consisting entirely of ordinary points. In a neighborhood of $t = t_1$ all of the points defined on the curve would also be points of S, since the solutions of the equations

$f = 0$ near the initial point p of the curve are all in S. The values of t defining points on the curve and in S would therefore have an upper bound $\tau \leqq t_2$ such that τ would define on the curve a boundary point of S. But this is impossible since all of the points of the curve are ordinary points.

If the functions f are known to be continuous and to have continuous derivatives in a region R, then it follows readily from what precedes that through any ordinary solution of the equations $f = 0$ interior to R there passes one and only one sheet of solutions having the property that the only boundary points of the sheet are boundary points of R, or interior points of R at which the functional determinant vanishes. If R is finite and closed and consists entirely of ordinary points for the functions f, then there can not be more than a finite number of points of the sheet on any ordinate x. Otherwise the points common to the ordinate and the sheet would have a point of condensation p, also in R. Since p is an ordinary point there can be at most one solution of the equations in a properly chosen neighborhood p_e.

It is interesting to determine a criterion which shall characterize a sheet which is at most single-valued on any ordinate. Such a criterion is derived in § 7 in connection with a theorem due originally to Schoenflies, and afterwards proved by Osgood. The proof of it involves the auxiliary notions described in § 6 and the following corollaries to the preceding theorem:

If the initial point of a continuous arc

$$(C_x) \qquad x_i = x_i(t) \qquad (i = 1, 2, \cdots, m; t' \leqq t \leqq t'')$$

in the x-space is the projection of a solution $p'(x'; y')$ of the equations $f = 0$ which is an ordinary point for the functions f, then there is associated with the arc C_x one and only one continuous curve

$$(C_{xy}) \quad x_i = x_i(t), \quad y_\alpha = y_\alpha(t) \qquad (i = 1, 2, \cdots, m; \alpha = 1, 2, \cdots, n)$$

passing through $(x'; y')$ for $t = t'$, with the properties:

1) *all of its points are solutions of the equations $f = 0$ and ordinary points of the functions f;*

2) *it is defined either over the whole interval $t' \leqq t \leqq t''$, or else on an interval $t' \leqq t < \tau$ $(\leqq t'')$ such that as t approaches τ the only limit points of the curve C'_{xy} are at infinity or are exceptional points of the functions f.*

The truth of this statement is readily deduced from the considerations which precede, or by the following argument. The fundamental theorem of § 1 can be applied at the point $(x'; y')$. If the arc C'_x is entirely within the region x_δ' then the existence and uniqueness of the curve C'_{xy} is evident. In any case there will be some intervals $t' \leqq t \leqq t_1$ in which curves C'_{xy} are defined having all the properties described in the theorem except possibly 2). Suppose that τ is the upper bound of the end values t_1 for such intervals. Then there is a curve C'_{xy} well defined in the interval $t' \leqq t < \tau$, and no limit point $(\alpha ; \beta)$ of the curve as t approaches τ can be a finite ordinary point for the functions f. For if $(\alpha ; \beta)$ were such a point, it would also satisfy the equations $f = 0$, on account of the continuity of the functions f, and the theorem of § 1 could again be applied at $(\alpha ; \beta)$. A curve C'_{xy} with all the properties of the theorem, except possibly 2), could then be defined over an interval including the interval $t' \leqq t < \tau$ in its interior, which contradicts the assumption that τ is the upper bound of such intervals.

There could not be two curves C'_{xy} associated with the projection C'_x, having the properties described in the theorem, and having distinct points $(x; y')$ and $(x; y'')$ corresponding to the same value t_2. For if so, there would be an interval $t_3 < t \leqq t_2$ in which the curves would be distinct while at $t = t_3$ they coincide. This is, however, impossible since in a neighborhood of the point corresponding to t_3 there can be but one solution of the equations $f = 0$ corresponding to a given set of values x.

Suppose that a continuum X of points $(x_1, x_2, \cdots, x_m)$ contains no projection of a boundary point of a sheet S of solutions of the equations $f = 0$, and no point near which the sheet becomes infinite. Then if X contains the projection of a point on the sheet every other point of X will also be such a projection. On the other hand, if X

contains a point which is not a projection of any point of the sheet, then no point of X can be a projection of a point of S.

These statements follow readily with the help of the last theorem. For suppose that X contains the projection x' of an interior point $(x'; y')$ of a sheet of solutions of the equations $f = 0$, and let x'' be any other point of X. Since X is a continuum there exists a continuous arc C_x entirely interior to X joining x' and x'', and the corresponding continuation curve C_{xy} must be defined over the whole of the arc C_x. Hence x'' is also the projection of a point of the sheet of solutions through $(x'; y')$. The rest of the theorem follows at once.

If the curve C_{xy} in the last theorem but one is defined over the whole arc C_x, and has initial and end points p' and p'', respectively, then there always exists a positive constant ρ such that any curve Γ_x lying in the ρ-neighborhood of the curve C_x and joining x' to x'', has a unique continuation curve Γ_{xy} also joining p' and p''.

The curve

$$(\Gamma_x) \qquad x = \xi_i(u) \qquad (i = 1, 2, \cdots, m;\ u' \leqq u \leqq u'')$$

is said to lie in the ρ-neighborhood of C_x if there exists a continuous function

$$(14) \qquad t = t(u) \qquad (u' \leqq u \leqq u'')$$

taking the values t', t'' at the ends of the u-interval, and such that the point α on Γ_x, defined by any value of u, lies in the neighborhood a_ρ of the corresponding point a of C_x determined by the relation (14).

It is possible to choose two constants, ϵ and $\delta \leqq \epsilon$, so that the neighborhoods p_ϵ and a_δ have the properties described in the theorem of § 1 uniformly for every point $p(a, b)$ on the arc C_{xy}. If not, there would be a sequence of points p_k on C_{xy} with a limit point π, for which the largest possible constants ϵ_k have the limit zero. But for the point π there is an effective constant $\epsilon > 0$, and the constants ϵ_k could not therefore decrease indefinitely in size. A similar argument shows the existence of the constant δ.

Suppose now that the interval $u' \leqq u \leqq u''$ is divided by values u_k $(k = 1, 2, \cdots, \nu)$ into sub-intervals so small that the points of any arc $\alpha_{k-1}\alpha_k$, corresponding on Γ_x to the values $u_{k-1} \leqq u \leqq u_k$, all lie in the $\frac{1}{2}\delta$-neighborhood of the point α_{k-1}, and further so small that the same is true with respect to the point a_{k-1} of the arc $a_{k-1}a_k$ of C_x corresponding to $\alpha_{k-1}\alpha_k$ by means of the relation (14). The constant ρ is supposed to have

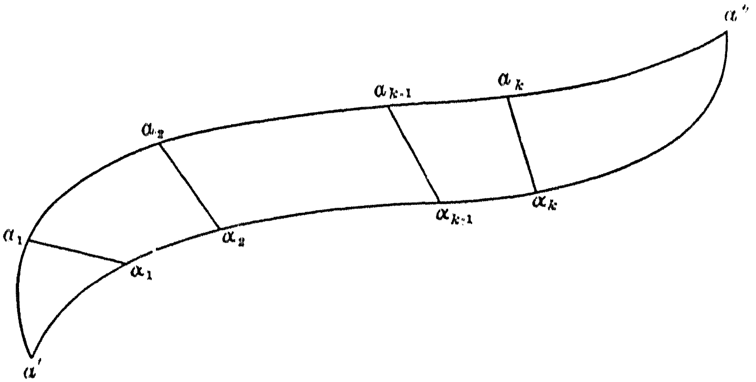

FIG. 1.

been chosen equal to $\frac{1}{2}\delta$, so that the curve Γ lies in the $\frac{1}{2}\delta$-neighborhood of C. Then the four-sided closed curve formed by the two straight lines $a_{k-1}\alpha_{k-1}$ and $a_k\alpha_k$, and the two arcs $a_{k-1}a_k$ and $\alpha_{k-1}\alpha_k$, lies entirely within the δ-neighborhood of the point a_{k-1}. The two continuation curves in the xy-space, starting with the point p_{k-1} on C'_{xy} and having as projections the arcs $a_{k-1}a_k\alpha_k$ and $a_{k-1}\alpha_{k-1}\alpha_k$, respectively, lead to the same point π_k corresponding to the point α_k in the x-space.

It is possible to argue, then, that the point π_1 on the continuation curve of the arc $a'\alpha_1$ is the same as that of the continuation curve for $a'a_1\alpha_1$, since the arcs $a'\alpha_1$ and $a'a_1\alpha_1$ lie entirely within the δ-neighborhood of the point a_1. Similarly, the point π_2 for the arc $a'\alpha_2$ is the same as that for the continuation curve along $a'a_2\alpha_2$. And finally the point π'' must coincide with p'', provided always that the initial points π' and p' of the continuation curves are the same.

In particular if the curve C_{xy} is defined over the whole arc C_x, as described above, then there exists a polygon in the x-space joining a' and a'' in the ρ-neighborhood of C_x, and along which there is a continuation curve in S also joining p' and p''. The polygon can be so chosen that no two adjacent sides have more than an end point in common.

To show this, let the interval $t' \leqq t \leqq t''$ be divided in any way by means of points of division $t', t_2, t_3, \cdots, t_\nu, t''$, and let the corresponding points on the curve C_{xy} be $(x'; y')$, $(\xi''; \eta'')$, $\cdots$, $(\xi^{(\nu)}; \eta^{(\nu)})$, $(x''; y'')$. The straight line $\xi^{(k)}\xi^{(k+1)}$ has the equations

$$x_i = \xi_i^{(k)} + \frac{t - t_k}{t_{k+1} - t_k}\left(\xi_i^{(k+1)} - \xi_i^{(k)}\right) \quad (i = 1, 2, \cdots, m).$$

Since the functions defining C_x are continuous, and therefore uniformly continuous, in $t' \leqq t \leqq t''$, it is possible to take the points of division $t', t_2, t_3, \cdots, t_\nu, t''$ so close together that the differences $x - \xi^{(k)}$, for any point x on the arc $\xi^{(k)}\xi^{(k+1)}$ of C_x, are uniformly less than an arbitrarily assigned positive constant δ; and the preceding theorem shows that the curve C_{xy} and the continuation curve along the polygon both lead from p' to p''.

If the sides $\xi^{(k)}\xi^{(k+1)}$ and $\xi^{(k+1)}\xi^{(k+2)}$ have more than the point $\xi^{(k+1)}$ in common, then one of the two would be included entirely within the other, and the continuation curve along $\xi^{(k)}\xi^{(k+2)}$ would have the same end points as that along the two successive sides. Therefore, by replacing adjacent sides by a single one whenever the two have more than one end point in common, a polygon as described in the theorem can be found.

§ 6. Auxiliary Theorems and Definitions

In this section it is proposed to record some theorems which will be of service later, especially in the proofs of the theorems of § 7. In the first place let it be agreed that a regular curve in the plane shall mean one which is continuous and has a well-defined tangent at all except possibly a finite number of points.

at each of which, however, the slope of the tangent approaches definite limits as the point is approached from either side. Analytically this means that the functions

$$x = x(t), \qquad y = y(t) \qquad\qquad (t' \leqq t \leqq t'')$$

defining a regular curve are continuous in the whole interval $t' \leqq t \leqq t''$, that they are differentiable and satisfy the inequality

$$(15) \qquad (dx/dt)^2 + (dy/dt)^2 \neq 0$$

at all except possibly a finite number of values of t. At an exceptional value $t = \tau$, where the derivatives are not well defined or where the expression (15) vanishes, the angle φ defined by the equations

$$\cos\varphi = \frac{dx/dt}{\sqrt{(dx/dt)^2 + (dy/dt)^2}}, \quad \sin\varphi = \frac{dy/dt}{\sqrt{(dx/dt)^2 + (dy/dt)^2}}$$

has nevertheless a unique limit as t approaches τ on the right, and a unique limit as t approaches τ on the left. These two limits are not necessarily the same.

It is known that a simply closed regular curve C in an xy-plane divides the plane into two continua, an exterior and a finite interior.* Any two interior points can be joined by a regular curve every point of which is an interior point, and a similar statement holds for exterior points. Any continuous curve joining an interior and an exterior point must have on it at least one point of the curve C, and any point p on C can be joined with an interior point by a regular curve which has in common with C only the point p.

The interior of a simply closed regular curve

$$x = x(t), \qquad y = y(t) \qquad\qquad (t' \leqq t \leqq t'')$$

can be divided by a finite number of segments of straight lines into

* See for example Osgood, Lehrbuch der Funktionentheorie, Chapter V, §§ 4–6; Bliss, "A proof of the fundamental theorem of analysis situs," *Bulletin of the American Mathematical Society*, vol. 12 (1906), page 336.

regions each of which has a maximum diameter less than an arbitrarily assigned positive constant ϵ.*

Let the maximum and minimum values of y in the interval $t' \leqq t \leqq t''$ be y_1 and y_2, and let p_1 and p_2 be two points of C at which y has these values. It is desired to show that there is a segment $p'p''$ of the horizontal line $y = (y_1 + y_2)/2$ which forms with C two simply closed regular curves, $p'p_1p''p'$ and $p'p_2p''p'$, each containing one of the points p_1 and p_2.

The points p_1 and p_2 can be joined by a regular curve D which, except at its end points, is interior to C. Two arcs of D adjoining p_1 and p_2, can be marked off in such a way that they do not cut the line $y = (y_1 + y_2)/2$. The remaining arc D' of D is entirely interior to C and can be replaced by a continuous polygon D'' with a finite number of sides, having the same end points and consisting also of interior points of C only. Any side of D'' which has an end point in common with the line $y = (y_1 + y_2)/2$ may be rotated slightly about its other end point, and in this way it may be brought about that D'' has only interior points of its sides on the line $y = (y_1 + y_2)/2$, and actually crosses the line wherever they have a point in common.

The polygon D'' must intersect $y = (y_1 + y_2)/2$ at least once, say at a point p, since one end point of D'' is above and the other below this line. There will be a segment $p'p''$ of $y = (y_1 + y_2)/2$, containing p and such that p' and p'' are on the curve C while every other point of the segment is interior to C. There can be only a finite number of such segments $p'p''$ containing points of D'', since D'' has at most a finite number of intersections with the horizontal line. There must be at least one segment on which D'' has an odd number of intersection points, since otherwise both end points of D'' would be on the same side of $y = (y_1 + y_2)/2$. If $p'p''$ is such a segment, then it forms with C two simply closed regular curves $p'p_1p''p'$ and $p'p_2p''p'$, one of which contains p_1 and the other p_2. For after its last intersection with $p'p''$ the polygon D'' and hence p_2 is entirely exterior to the curve $p'p_1p''p'$.

* For a similar theorem see Osgood, loc. cit., Chapter V, § 9.

For the moment that part of a curve which does not lie in a horizontal line may be called the effective arc of the curve, in view of the fact that the altitude of the curve can not be more than one half the length of this so-called effective part. If the altitude of any curve is $\geqq \epsilon$, the effective length of either of its two parts after subdivision by a horizontal line segment, as described above, will be $\leqq L - \epsilon$, where L is its effective length.

If the altitude $y_1 - y_2$ of C is greater than ϵ, then the effective arc of either $p'p_1p''p'$ or $p'p_2p''p'$ will be greater in length than ϵ, and the effective length of each will also be less than $L - \epsilon$, where L is the perimeter of C. If the curve $p'p_1p''p'$, for example, has still an altitude greater than ϵ, it may be subdivided by a horizontal segment as before, and the effective arcs of the two new curves so found will be less than $L - 2\epsilon$. By a continuation of this process the interior of C will be subdivided finally by curves whose effective lengths are less than 2ϵ and whose altitudes are therefore less than ϵ.

In a similar manner the regions so formed may be subdivided by vertical segments into others whose breadths are less than ϵ, and the theorem follows at once.

A set of points in an x_1x_2-plane is *connected* if any two of its points can be joined by a continuous arc whose points all belong to the set, and it is further said to be *simply connected* if every simply closed regular curve in it has an interior which also consists only of points of the set.

It is more difficult to set down a satisfactory definition of simple connectivity for sets of points in an m-dimensional space. In the following section of these lectures, however, a special type of simple connectivity is needed which may be defined by means of some simple auxiliary conceptions.

A *normal subspace of two dimensions* in a region X of points $(x_1, x_2, \cdots, x_m)$ is a totality of points defined by equations of the form

$$x_i = \varphi_i(u_1, u_2) \qquad (i = 1, 2, \cdots, m),$$

where

1) the values (u_1, u_2) range over a simply connected region U;

2) no two distinct sets of values u define the same point x;

3) the functions φ are continuous and have continuous first derivatives in U;

4) the determinants of the second order of the matrix of derivatives $||\partial\varphi_i/\partial u_k||$ $(i = 1, 2, \cdots, m;\ k = 1, 2)$ do not all vanish simultaneously at any point of U.

A simply connected region in two dimensions is defined above, and a connected region X in a space of points $(x_1, x_2, \cdots, x_m)$ has a definition quite similar to that for two dimensions. In order to specify conveniently the properties of a region X which is simply connected, the term *elementary curve* will also be used. By an elementary curve in X is meant a simply closed continuous curve which either lies in a normal subspace of two dimensions entirely in the interior of X, or else is such that in every neighborhood of it there is a simply closed continuous curve having this property. It is thus seen that while an elementary curve may not itself be imbedded in one of the two-dimensional normal subspaces interior to X, it can nevertheless be approximated as closely as may be desired by one which does. The word neighborhood is here used in the sense described in connection with the fourth theorem of § 5 (see page 26).

If a region X is connected, then any simply closed continuous curve in its interior may be developed into two such curves by an auxiliary arc joining two of its points, and the process of development may be continued on the two arcs so formed.

*If a region X is such that any simply closed continuous curve in its interior is an elementary curve, or may be developed into a number of elementary curves by means of auxiliary arcs, as just described, then X is said to be simply connected.**

* For a discussion of the connectivity of higher spaces, see Picard and Simart, Théorie des Fonctions algébriques de deux Variables indépendantes, Chapitre II, in particular §§ 11 ff. If every simply closed continuous curve interior to R lies in a normal subspace of two dimensions interior to R, one sees intuitively that a second neighboring subspace of the same kind can be passed through the curve. The closed two-dimensional subspace so formed is

§ 7. A Criterion that a Sheet of Solutions be Single-valued

Consider in the first place a set of equations

$$f_\alpha(x_1, x_2;\ y_1, y_2, \cdots, y_n) = 0 \quad (\alpha = 1, 2, \cdots, n) \tag{16}$$

in which there are but two independent variables x.

If a connected sheet S of solutions of equations (16) *consists only of ordinary points of the functions f, and furthermore has a simply connected projection X in the x_1x_2-plane such that no interior point of X is either a point where S becomes infinite or the projection of a boundary point of S, then the sheet S is single-valued over the interior of X.*

Suppose, in contradiction to the theorem, that over any interior point of X there were two points, p' and p'', of the sheet. Since S is connected there would be a continuous curve

$$(C_{xy}) \quad x_1 = x_1(t), \quad x_2 = x_2(t), \quad y_\alpha = y_\alpha(t)$$
$$(t' \leqq t \leqq t'';\ \alpha = 1, 2, \cdots, n)$$

consisting entirely of interior points of the sheet and joining p' with p'' in the space $(x; y)$. The projection

$$(C_x) \qquad x_1 = x_1(t), \qquad x_2 = x_2(t) \qquad (t' \leqq t \leqq t'')$$

of this curve would necessarily be a closed curve in the x_1x_2-plane, and by the second theorem of § 5 the arc C_{xy} is the only one associated with C_x in the sheet S and having the initial point p'.

The curve C_x may be simply closed and regular; but if it is not, there will nevertheless be a curve in the region X having these properties, and for which the continuation curve analogous to C_{xy} is not closed. For, in the first place, from § 5 it is seen that the curve C_x may be supposed to be a polygon no two adjacent sides of which have more than an end point in common, provided that it is desired only to secure a continuous curve in

separated into two parts by the curve, and hence the number which Picard and Simart designate by p_1 is equal to unity for a simply connected region of the kind defined in the text above.

4

the sheet passing from p' to p''. Let the corners of this polygon in the x-plane be denoted by $\xi_1, \xi_2, \cdots, \xi_\nu$, where ξ is a symbol for a point (x_1, x_2). The side $\xi_\nu\xi_1$ touches $\xi_1\xi_2$ at its end point ξ_1, and it can be argued therefore that there will be some first side $\xi_\lambda\xi_{\lambda+1}$ which touches some one of the preceding sides elsewhere than at its initial point ξ_λ. Let the side so touched by $\xi_\lambda\xi_{\lambda+1}$ be $\xi_\kappa\xi_{\kappa+1}$, where $\kappa + 1$ is necessarily less than λ, and let the first point of $\xi_\lambda\xi_{\lambda+1}$ which lies on $\xi_\kappa\xi_{\kappa+1}$ be ξ. If the portion of the curve C_{xy} which corresponds to the polygon

$$(17) \qquad \xi, \quad \xi_{\kappa+1}, \quad \xi_{\kappa+2}, \quad \cdots, \quad \xi_\lambda, \quad \xi$$

is not closed, then the polygon (17) itself is a simply closed curve in X of the kind desired above, that is, one along which there exists a continuation curve in the xy-space whose end points are different.

If the portion of C_{xy} which corresponds to (17) is closed, then that part of C_{xy} which belongs to the polygon

$$(18) \qquad \xi_1, \ \xi_2, \ \cdots, \ \xi_\kappa, \ \xi, \ \xi_{\lambda+1}, \ \cdots, \ \xi_\nu, \ \xi_1$$

is also continuous and leads from p' to p''. Since $\kappa + 1 < \lambda$ the side $\xi_{\kappa+1}\xi_{\kappa+2}$ at least is missing in (18), and the number of sides is at least one less than that of the original polygon. By an alteration of the kind suggested in the proof of the last theorem of § 5, which also reduces the number of sides, it can be brought about, if not already true, that the polygon (18) still has no two adjacent sides with more than an end point in common.

By continuing this process one must come at some stage to a simply closed regular curve in the x-plane with a corresponding continuation curve in the xy-space which is not closed. In order not to complicate the notation too much it may be supposed that the curve C_x itself is such a curve. Every point of C_x is an interior point of the region X since the corresponding point of C_{xy} is an interior point of the sheet S. The interior of C_x is therefore also composed entirely of interior points of X, since X is simply connected. If the interior of C_x is subdivided into

two parts by a segment of a straight line, as described in the preceding section, the dividing segment will also have a continuation curve on the sheet S throughout its entire length, by the second theorem of § 5. For its initial point on the curve C_x corresponds to an interior point of the sheet S and, by the hypothesis of the theorem which is to be proved, none of its points can be a point where S becomes infinite or can correspond to a boundary point of S. Hence one of the simply closed curves formed by the curve C_x and the dividing segment is a curve retaining the property that it has a continuation curve on the sheet S which is not closed. Suppose that C_x' is this curve. By continuing the process a sequence of curves $\{C_x^{(k)}\}$, with diameters approaching zero, can be found, each lying in the interior of C_x and having an unclosed continuation curve $C_{xy}^{(k)}$ on S.

If a point $p^{(k)}$ is selected arbitrarily on the curve $C_{xy}^{(k)}$, the sequence $\{p^{(k)}\}$ $(k=1, 2, \cdots, \infty)$ will have a finite point of condensation $\pi(\alpha; \beta)$ in the xy-space which is an interior point of the sheet S. For the projections $x^{(k)}$ of the points $p^{(k)}$ all lie in the interior of C_x and hence must have a point of condensation α. Furthermore the points of the sequence $p^{(k)}$ whose projections are in the neighborhood of α can not become infinite or approach a boundary point of the sheet, since α is interior to X. They must therefore have at least one limit point π which is an interior point of the sheet, and with which there are associated two neighborhoods π_ϵ and α_δ by the principal theorem of § 1. Some of the points $p^{(k)}$ lie in π_ϵ, and have corresponding curves $C_x^{(k)}$ in α_δ. For such points the continuation curves $C_{xy}^{(k)}$ also lie in π_ϵ and can not be unclosed, since to any point x in α_δ there corresponds in π_ϵ at most one solution of the equations $f=0$. The original assumption that S is multiple-valued in the interior of X is therefore contradicted.

The theorem remains true for any system of equations of the form

$$(19)\quad f_\alpha(x_1, x_2, \cdots, x_m;\ y_1, y_2, \cdots, y_n) = 0 \quad (\alpha = 1, 2, \cdots, n).$$

With a continuous curve C joining (u_1', u_2') to an arbitrarily chosen point (u_1, u_2) of U there is always associated a continuation curve of solutions of the equations (21), having the initial point p_u' and defined throughout the whole of C, since any such curve defines a curve in the x-space interior to X along the whole of which there is a corresponding continuation curve for the equations (19) in the sheet S. Hence there is a unique sheet S_u of solutions of the equations (21) whose projection in the u_1u_2-space is U; and no interior point of U is a point where the sheet becomes infinite or corresponds to a boundary point of the sheet, since the same is true of S with respect to X. The preceding argument can therefore be applied to show that the sheet S_u is single-valued over the region U, and the existence of the curve C_{uy} with the distinct end points p_u' and p_u'' is contradicted. Hence C_{xy} can not have distinct end points p' and p'', and the theorem last stated is proved.

§ 8. Transformations of n Variables and a Modification of a Theorem of Schoenflies

It is interesting to deduce by means of the preceding theorems some conclusions concerning a system of equations of the form

$$(22)\quad f_\alpha(x; y) = x_\alpha - \psi_\alpha(y_1, y_2, \cdots, y_n) = 0 \quad (\alpha = 1, 2, \cdots, n).$$

The functions ψ are once for all assumed to be single-valued, continuous, and to have continuous first derivatives in a continuum Y in which the functional determinant

$$D = \partial(\psi_1, \psi_2, \cdots, \psi_n)/\partial(y_1, y_2, \cdots, y_n)$$

is different from zero. By a continuum is meant a set of points consisting only of interior points any two of which can be connected by a continuous curve lying entirely within the set. The boundary points of Y will be denoted by B, and X will represent the set of points in the x-space which corresponds to Y by means of the equations (22).

Any sequence $\{y^{(k)}\}$ of points $(y_1^{(k)}, y_2^{(k)}, \cdots, y_n^{(k)})$ $(k = 1, 2, \cdots)$ in Y, which approaches infinity or has a point of B as limit point, defines a corresponding sequence of points $\{x^{(k)}\}$ in X. The set of points of condensation for such sequences $\{x^{(k)}\}$ will be denoted by A.

The totality of solutions of the equations (22) *corresponding to points of the continuum Y form a single connected sheet S whose only boundary points have projections x and y in the sets A and B, respectively.*

For suppose that $(x'; y')$ is a first solution and $(x''; y'')$ any other. The points y' and y'' can be joined by a continuous curve interior to Y

$$y_\alpha = y_\alpha(t) \qquad (\alpha = 1, 2, \cdots, n;\ t' \leq t \leq t''),$$

and the corresponding curve

$$x_\alpha = x_\alpha(t), \qquad y_\alpha = y_\alpha(t),$$

defined by equations (22), is a curve interior to the sheet S and joining $(x'; y')$ to $(x''; y'')$, so that S is evidently connected. Any boundary point $(\alpha; \beta)$ of S must be the limit of a sequence of points $p^{(k)}$ for which the projections y are in Y. The limit β of the sequence $y^{(k)}$ can not be in Y, since then $(\alpha; \beta)$, by the theorem of § 1, would be an interior point of S. Hence β must be in B and α in A.

One may say further that if $p^{(k)}$ is a sequence of points $(x^{(k)}; y^{(k)})$ in S for which the sequence $x^{(k)}$ approaches infinity, then the only finite points of condensation possible for the sequence $y^{(k)}$ are in B. The statement is true when x and y are interchanged, on account of the definition above of the set A.

If the points of the set A are distinct from those of the image X of Y, then X is a single continuum whose only boundary points are points of A.

To prove this, consider an arbitrarily chosen point y' of Y. None of the points in a suitably chosen neighborhood of the corresponding values x' are points of A, since by the fundamental

theorem of § 1 all such points correspond by means of equations (22) to points of Y, and are therefore points of X. Consider now the continuum $\bar{X}$ consisting of all points x which can be joined to x' by continuous curves containing no points of A, a continuum to which the neighborhood of x' certainly belongs, as has just been shown.

All the points of X are in the continuum $\bar{X}$, since the solutions of equations (22) corresponding to points of Y form a single connected sheet S. The curve in S joining $(x'; y')$ with any other point $(x''; y'')$ of the sheet has therefore a projection in the x-space joining x' with x'' and containing no points of the set A.

All of the points of $\bar{X}$ are points of X. For any set of values x in $\bar{X}$ can be joined to x' by a continuous curve C_x lying entirely in $\bar{X}$ and containing therefore no points of A. By the second theorem of § 5 the corresponding continuation curve C_{xy} must extend along the entire arc C_x, since otherwise the values of y for points on C_{xy} would approach infinity or else have a limit point on the boundary B of Y, and some point of C_x would in that case necessarily be a point of A. It follows that x, like x', is the image of some point y in Y.

From the initial theorem of the last section, for the case when there are more than two variables, it follows that

If A is distinct from X, and X is simply connected in the sense of §6, then the sheet S is single-valued. In other words the continuum Y is tranformed in a one-to-one way into a continuum X by means of the equations (22), and the functions

$$y_\alpha = y_\alpha(x_1, x_2, \cdots, x_n) \qquad (\alpha = 1, 2, \cdots, n) \tag{23}$$

so defined over X are single-valued, continuous, and have continuous first derivatives.

The character of the functions (23) near any point of X follows at once from the theorem of § 1.

Let it be supposed that the set of points A divides the x-space into exactly two continua X, Ξ such that every point of A is a bound-

ary point for each of them, and suppose furthermore that there is a particular point ξ in Ξ which does not correspond by means of the equations (22) *to any point of Y. Then the image X of Y is distinct from A and coincides with X. If X is simply connected the other conclusions of the last theorem follow at once.*

In the first place it can be shown that if any point ξ' of Ξ corresponds to a point of Y then every other point ξ'' of Ξ would also have this property. For ξ' and ξ'' can be joined by a continuous curve

$$x_\alpha = x_\alpha(t) \qquad (\alpha = 1, 2, \cdots, n;\ t' \leqq t \leqq t'')$$

entirely interior to Ξ. The corresponding continuation curve

$$x_\alpha = x_\alpha(t), \qquad y_\alpha = y_\alpha(t)$$

of solutions of equations (22) must be defined along the whole of the interval $t' \leqq t \leqq t''$, since otherwise as t approached any upper bound τ of the values t which could be reached by continuation, the corresponding points y of the curve would have to approach infinity or else have a point of condensation on the boundary of Y. But this is impossible, since for a sequence of points x corresponding to a sequence of points in Y approaching infinity or a boundary point of Y, the only limiting points possible are at infinity or else in the set A. It follows at once, on account of the hypothesis of the theorem, that no point of Ξ can correspond to a point of Y, and neither can any point of A, since in any neighborhood of such a point of A there are points of Ξ which in that case would also correspond to values y in Y. The image of the region Y in the x-space is a single continuum whose only boundary points are points of A. According to the preceding argument it cannot be Ξ and it must therefore be X.

A modification of a theorem of Schoenflies can be deduced readily from the results which precede. The theorem has to do with a pair of equations of the form

$$(24) \qquad x_1 = \psi_1(y_1, y_2), \qquad x_2 = \psi_2(y_1, y_2)$$

in which the functions ψ are single-valued, continuous, and have continuous derivatives on a simply closed regular curve B of the y-plane and in the interior Y of B. The functional determinant $D = \partial(\psi_1, \psi_2)/\partial(y_1, y_2)$ is supposed to be different from zero in Y.

If the curve A in the x-plane formed by transforming the simply closed regular curve B in the y-plane, by means of the equations (24), *is distinct from the image X of the interior Y of B, then X is a simply connected continuum whose only boundary points are points of A, and the correspondence defined between X and Y is one-to-one. The single valued functions*

$$(25) \qquad y_1 = y_1(x_1, x_2), \qquad y_2 = y_2(x_1, x_2),$$

*so determined in the region X, are continuous and have continuous first derivatives.**

From the preceding theorems of this section it follows that the complete image X of Y is a single finite continuum whose only boundary points are points of A. It remains to show that X is simply connected and that the correspondence between X and Y is one-to-one.

If any simply closed regular curve C_x is drawn in X, its interior must consist entirely of points of X. Otherwise there would necessarily be a boundary point of X, a point of the curve A, interior to C_x, and there would also be points of A exterior to C_x since X is finite. Hence there would necessarily be a point of the continuous curve A on C_x itself, which contradicts the assumption that A and X are distinct. It follows at once from the first paragraphs of § 7 and the simple connectivity of X just proved, that only one point y in Y corresponds to a given x in X, and by the theorem of § 1 it may be seen that the functions

* Schoenflies assumed only the continuity of the functions ψ_1, ψ_2, adding, however, that the correspondence defined between the regions X and Y of the two planes is to be one-to-one. In the theorem here proved ψ_1 and ψ_2 are subjected to further continuity restrictions, but the correspondence is proved to be unique. See Schoenflies, "Ueber einen Satz der Analysis Situs," *Göttinger Nachrichten* (1899), page 282. The theorem was later proved by Osgood and Bernstein in the same journal (1900), pages 94 and 98, respectively.

(25) have the continuity properties described in the theorem in the neighborhood of any particular point x.

Another theorem, slightly different in form, may be stated as follows:

If the images of the points of the simply closed regular curve B in the y-plane all lie on a simply closed regular curve A in the x-plane, then the equations (24) *define a one-to-one correspondence between the interior X of A and the interior Y of B, and the functions* (25) *so defined have the same continuity properties as before.*

In this case it can first be shown that the image x' of any point y' in Y must be distinct from A, and the rest of the proof is the same as before. For, if x' were a point of A, every point of a properly chosen neighborhood of x' would also be the image of a point of Y, since at $(x'; y')$ the functional determinant of equations (24) does not vanish. It would follow then, by continuation, that every point exterior to the curve A would also be the image of a point of Y, which is impossible since the functions ψ are finite. The continuum X is therefore identical with the interior of A, by the preceding theorems, and the correspondence between X and Y is one-to-one.

An example applying some of the theorems of §§ 5, 8 is given at the end of § 14.

CHAPTER II

SINGULAR POINTS OF IMPLICIT FUNCTIONS

The theorems which have been developed in the preceding pages of these lectures have to do with the behavior of implicit functions at ordinary points, or in regions which have no singular points in their interiors. For singular points where the functional determinant vanishes the theory is much more complicated, and no methods which can be comprehensively applied have so far been developed. There are, however, many special cases in widely different fields which have been studied with success, and it may not be out of place to glance at a few of them before proceeding to the further theorems with which these pages are primarily concerned.

Perhaps the most complete single theory which has been developed is that which has to do with the singularities of an algebraic function y of x determined by an equation of the form

$$P(x, y) = 0, \tag{1}$$

where P is an irreducible polynomial in the two variables x and y. Suppose for convenience that the singular point to be considered is at the origin, and that the polynomial $P(0, y)$ has a lowest term of degree n in y. Then it is known that for each value of x in a sufficiently small neighborhood of $x = 0$, there exist exactly n solutions y of equation (1) in the neighborhood of $y = 0$, and the values of these solutions are given by k cycles of the form

$$y = a_j x^{\mu_j/p_j} + a_j' x^{\mu_j'/p_j} + \cdots \quad (j = 1, 2, \cdots, k), \tag{2}$$

where the numbers μ, p are positive integers satisfying the relations

$$\mu_j < \mu_j' < \mu_j'' < \cdots, \qquad p_1 + p_2 + \cdots + p_k = n.$$

The series is one member of the cycle; the others are found by replacing x^{1/p_j} by $\omega^{\nu}x^{1/p_j}$ ($\nu = 1, 2, \cdots, p_j - 1$), where ω is a primitive p_j-th root of unity. The number p_j has no factor in common with the exponents $\mu_j, \mu_j', \cdots$. Otherwise the expansion would be in terms of a root of x of lower order than p_j. Thus there are in all n series in fractional powers of x which define the roots of the algebraic equation in the neighborhood of the origin. The coefficients of the series may be computed by means of the well-known Newton polygon,* or by methods due to Hamburger† and Brill.‡ If the substitution $x = t^{p_j}$ is made in the series (2), the points (x, y) which it defines may be expressed in the parametric representation

$$x = t^{p_j}, \qquad y = t^{\mu_j}\{\alpha_j + \alpha_j' t^{\mu_j' - \mu_j} + \cdots\} \qquad (j = 1, 2, \cdots, k).$$

All the solutions of the equation (1) in the neighborhood of the origin evidently belong to a finite number of such branches.

With the help of the preparation theorem of Weierstrass, which is to be studied in the following pages, results similar to those just given may be proved for the solutions of an equation $F(x, y) = 0$ in the vicinity of any point where F is analytic.

The singularities of a surface

$$F(x, y, z) = 0$$

at a point where the function F is analytic have also been extensively studied. The points of the surface in the neighborhood of a singular point are determined by means of a finite number of expansions of the form

$$x = P(u, v), \qquad y = Q(u, v),$$

where P and Q are analytic in the parameters u and v.§

---

* See Appell and Goursat, Théorie des Fonctions algébriques, pp. 184 ff.

† Weierstrass, Werke, vol. 4, Kapitel 1.

‡ *Münchener Berichte*, vol. 21 (1891), p. 207.

§ See Black, "The parametric representation of the neighborhood of a singular point of an analytic surface," *Proceedings of the American Academy of Arts and Sciences*, vol. 37 (1902), p. 281.

In the calculus of variations the construction of "fields of extremals" in the plane requires the study of the real solutions of a system of equations of the form

$$(3) \qquad x = \varphi(t, a), \qquad y = \psi(t, a).$$

The extremals are the curves in the xy-plane defined by these equations for different values of a. Suppose that the parametric values

$$(4) \qquad t_0 \leq t \leq t_1, \qquad a = a_0$$

define an arc E which does not intersect itself and which consists entirely of points where the functional determinant

$$(5) \qquad \Delta(t, a) = \frac{\partial(\varphi, \psi)}{\partial(t, a)}$$

is different from zero. Then to any point (x, y) in a properly chosen neighborhood of E there corresponds but one solution (t, a) of equations (3), in the neighborhood of the values (4); and the functions

$$t = t(x, y), \qquad a = a(x, y)$$

so defined have continuity properties similar to those of φ and ψ themselves.* The neighborhood thus simply covered by the extremals (3) is the "field," and is perhaps the simplest example of the notion since it consists only of non-singular solutions of the equations (3).

When it is desired to find an arc C which minimizes an integral with respect to variations lying entirely on one side of C, a field of a different sort can be constructed.† The equations of the

The mathematical literature concerned with the singularities of a curve or surface, particularly their transformation into simpler types, is very large. The reader is referred to Pascal, Repertorium der höheren Mathematik, 2d edition, vol. 2, erste Hälfte, pp. 291 ff; and Encyclopädie der Mathematischen Wissenschaften, II B 2, p. 119, and III C 4, pp. 365 ff.

* Bolza, Vorlesungen über Variationsrechnung, pp. 249 ff.

† Bliss, "Sufficient conditions for a minimum with respect to one-sided variations," *Transactions of the American Mathematical Society*, vol. 5 (1904), p. 477; Bolza, "Existence proof for a field of extremals tangent to a given curve," ibid., vol. 8 (1907), p. 399.

extremals (3) can be taken so that for $t = 0$ they all intersect C and are tangent to it, and the equations

$$x = \varphi(0, a), \qquad y = \psi(0, a)$$

will then be the equations of C. If the curvatures of the two arcs at their point of contact are always different, then the extremal arcs E simply cover a portion of the plane N on one side of C and adjacent to it. In other words, the equations (3) define a one-to-one correspondence between the points of a region adjoining the axis $t = 0$ in the ta-plane, shown in the accompanying figure, and a certain neighborhood N on one side of the arc C.

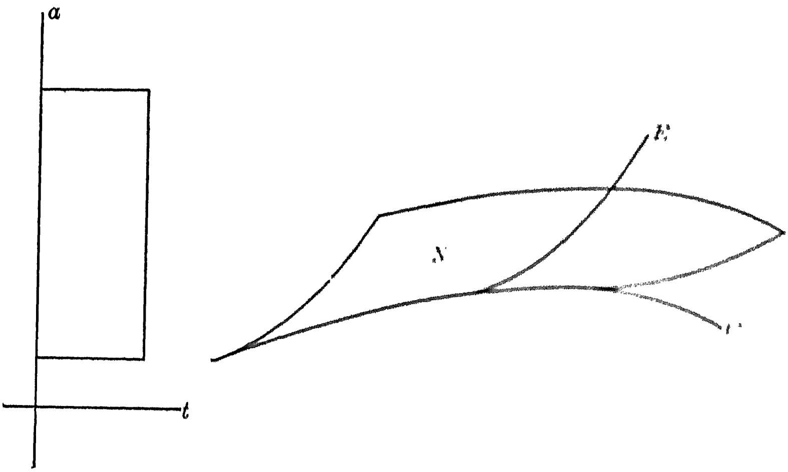

FIG. 2.

In the interior of the region N the functions $t(x, y)$, $a(x, y)$ have continuity properties similar to those of φ and ψ themselves. It is easy to see that this is a case in which the functional determinant (5) vanishes along the boundary $t = 0$ of the region to be transformed, since the curves C and E are always tangent.

In a paper published since these lectures were given, Dr. E. J. Miles* has considered the transformation defined by the equations

* "The absolute minimum of a definite integral in a special field," *Transactions of the American Mathematical Society*, vol. 13 (1912), pp. 37 ff.

(3) when the curve C to which the extremals E are tangent has a cusp, a situation corresponding to still another problem in the

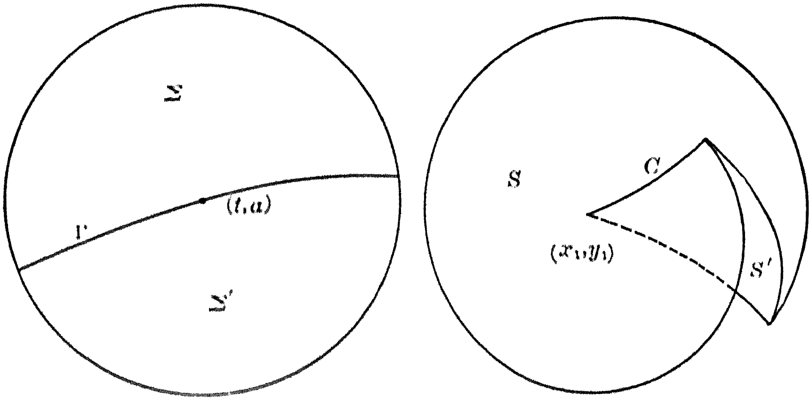

FIG. 3.

calculus of variations. In that case a point (t_1, a_1) and a curve Γ through it are transformed into a point (x_1, y_1) and a curve C as shown in the figure. One portion Σ of a neighborhood of (t_1, a_1) is then transformed in a one-to-one way into the leaf S, and the other portion Σ′ into the leaf S'. At any point in the interior of one of the leaves, the variables t and a are single-valued functions of x, y having continuity properties similar to those of φ and ψ. The transformation is singular along the curve Γ.

The three examples which have been just described are only a few of the many proofs for the existence of fields involving transformations with singular points which might be cited.* Nearly all of these have to do with singularities of transformations of the form

(6) $$x = \varphi(u, v), \qquad y = \psi(u, v),$$

* Bliss, "The construction of a field of extremals about a given point," *Bulletin of the American Mathematical Society*, vol. 13 (1906), p. 47; Mason and Bliss, "Fields of extremals in space," *Transactions of the American Mathematical Society*, vol. 11 (1910), p. 325; Bill, "The construction of a space field of extremals," *Bulletin of the American Mathematical Society*, vol. 15 (1908), p. 374; Szücs, "Sur l'extrémale qui joint deux points donnés," *Mathematische Annalen*, vol. 71 (1912), p. 380. The method used by Szücs is quite closely that of Mason and Bliss in the paper mentioned above.

or

$$x = \varphi(u, v, w), \qquad y = \psi(u, v, w), \qquad z = \chi(u, v, w),$$

which have been studied also in a series of papers of more recent date presented as dissertations for the degree of doctor of philosophy at Harvard University.* The methods which have been used in the different cases have differed widely, and it does not seem possible at present to formulate a theory which includes them all. It is the intention of the writer, however, to show in the following pages how the transformation theorems proved above in § 7 may be applied to throw much light on the nature of real transformations of the form (6) in the neighborhoods of singular points. In the section of the lectures immediately following this introduction a simple algebraic proof of the preparation theorem of Weierstrass is given, not depending upon the theory of functions of a complex variable. A generalization of it is given in a later section which, in what might be called the general case, enables one to describe the behavior of the solutions of a system of equations of the form

$$f_i(x_1, x_2, \cdots, x_m; y_1, y_2, \cdots, y_n) = 0 \quad (i = 1, 2, \cdots, n)$$

in the neighborhood of a point where the functional determinant

$$\frac{\partial(f_1, f_2, \cdots, f_n)}{\partial(y_1, y_2, \cdots, y_n)}$$

vanishes. For these equations the variables x and y are permitted to have complex values.†

* Urner, "Certain singularities of point transformations in space of three dimensions," *Transactions of the American Mathematical Society*, vol. 13 (1912), p. 233; Clements, "Implicit functions defined by equations with vanishing jacobian," to appear in the same journal. Dederick, in a paper entitled "The solutions of an equation in two real variables at a point where both the partial derivatives vanish," *Bulletin of the American Mathematical Society*, vol. 16 (1909), p. 174, has discussed the singularities of a curve of the form $F(x, y) = 0$ with the help of a sort of generalization of the Weierstrass preparation theorem for a function which is not necessarily analytic.

† The proof given in these pages for the last-mentioned theorem is for the case of two variables y. For n variables see the reference in the last footnote to § 13.

§ 9. The Preparation Theorem of Weierstrass

The theorem which is to be proved may be stated in the following form:

Let $f(x_1, x_2, \cdots, x_m, y)$ be a convergent series in the variables x, y, and such that the series $f(0, 0, \cdots, 0, y)$ begins with a term of degree n. Then f is factorable in the form

$$f(x_1, x_2, \cdots, x_m, y) = (y^n + a_1 y + \cdots + a_n)\varphi(x_1, x_2, \cdots, x_m, y),$$

where $a_1, a_2, \cdots, a_n$ are convergent power series in $x_1, x_2, \cdots, x_m$ which vanish for $x_1 = x_2 = \cdots = x_m = 0$, and φ is a power series in $x_1, x_2, \cdots, x_m, y$ which has a constant term different from zero.

In the *Bulletin de la Société Mathématique de France** Goursat has called attention to the fact that the proof which Weierstrass gave of this important theorem, as well as the later proofs which occur in the literature†, make use of the notions of the function theory, while the theorem itself is essentially of an algebraic character. In the paper referred to he has given an elegant and elementary proof of the theorem which is in outline as follows:

By means of the substitution

$$y^n = -a_1 y^{n-1} - a_2 y^{n-2} - \cdots - a_n$$

the series f can be reduced to a polynomial P of degree $n-1$ in y, whose n coefficients are convergent series in $a_1, a_2, \cdots, a_n, x_1, x_2, \cdots, x_m$. By the usual theorems of implicit function theory it is shown that the n equations found by putting these coefficients equal to zero have unique solutions for $a_1, a_2, \cdots, a_n$ as power series in $x_1, x_2, \cdots, x_m$, which vanish with $x_1, x_2, \cdots, x_m$. If the values so found are substituted in the formula

$$y^n = -a_1 y^{n-1} - a_2 y^{n-2} - \cdots - a_n + \mu$$

* "Démonstration élémentaire d'un théoreme de Weierstrass," vol. 36 (1908), p. 209.

† See, for example, Picard, Traité d'Analyse, vol. 2, p. 243; Goursat, Cours d'Analyse, vol. 2, p. 284.

and the series f again reduced, a polynomial P_1 of degree $n-1$ in y will be found whose coefficients are series in $x_1, x_2, \cdots, x_m, \mu$. On account of the way in which the functions $a_1, a_2, \cdots, a_n$ were determined, this polynomial P_1 has a factor μ, and hence f has a factor $(y^n + a_1y^{n-1} + \cdots + a_n)$.

Since the paper of Goursat appeared two further proofs of the theorem have been published, one by the writer* and the other by MacMillan,† each of which seems even more direct than that of Goursat. In the proof which follows use is made of the very concise and elegant method of MacMillan for determining the coefficients, while the rest of the proof is similar to that of the earlier paper of the writer cited above.

The theorem may be stated in a different form as follows:

Suppose that $f(x_1, x_2, \cdots, x_m, y)$ *is a series with literal coefficients such that* $f(0, 0, \cdots, 0, y)$ *begins with the term* a_0y^n. *Then there is one and but one series* $b(x_1, x_2, \cdots, x_m, y)$ *which satisfies formally the relation*

$$bf = p, \tag{7}$$

where p *is a polynomial*

$$p = a_0y^n + a_1y^{n-1} + \cdots + a_n$$

whose coefficients $a_k(x_1, x_2, \cdots, x_m)(k = 1, 2, \cdots, n)$ *are series vanishing with the* x*'s.*

Each of the coefficients in b *and the* a*'s is a rational function of a finite number of the coefficients of* f *with denominator a power of* a_0, *and the constant term in* b *is unity.*

If the coefficients in f *are chosen numerically so that* f *converges and* $a_0 \neq 0$, *then the series* b *and* a_k $(k = 1, 2, \cdots, n)$ *also converge.*

The functions f, b, p may be written in the forms

$$\begin{aligned} f &= a_0y^n - y^{n+1}f_0 - f_1 - f_2 - \cdots, \\ b &= b_0 + b_1 + b_2 + \cdots, \\ p &= a_0y^n - p_1 - p_2 - \cdots, \end{aligned} \tag{8}$$

* *Bulletin of the American Mathematical Society*, vol. 16 (1910), p. 356.
† Ibid., vol. 17 (1910), p. 116.

where f_k, b_k, p_k are homogeneous expressions of degree k in $x_1, x_2, \cdots, x_m$ with coefficients which are series in y. It is desired to determine b so that the identity (7) holds, and so that the expressions p_k have coefficients which contain y only to the degree $n-1$.

By substituting the expressions (8) in the identity (7) and equating terms of the same degree in the x's, it follows that

$$\begin{aligned}
&b_0(a_0 - yf_0)y^n = a_0y^n,\\
&b_1(a_0 - yf_0)y^n = b_0f_1 - p_1,\\
&b_2(a_0 - yf_0)y^n = b_0f_2 + b_1f_1 - p_2,\\
&\quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot,\\
&b_k(a_0 - yf_0)y^n = b_0f_k + b_1f_{k-1} + \cdots + b_{k-2}f_2 + b_{k-1}f_1 - p_k,\\
&\quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot.
\end{aligned}$$

These equations are to be identities in x and y. The first one determines b_0 uniquely with constant term unity, and furthermore so that each coefficient is a quotient, in fact a polynomial with positive integral coefficients in a finite number of the coefficients of f, divided by a power of a_0. In the second equation p_1 must be chosen equal to the terms of b_0f_1 which contain y to the degree $n-1$ or less, after which b_1 is uniquely determined. Similarly in the kth equation p_k must first be chosen to cancel the terms on the right of degree $n-1$ or less in y, and then b_k is unique.

It only remains to show that the series b and a_k are convergent in any numerical case for which f converges. There is no loss of generality in assuming that the series f converges in the domain

$$|x_i| \leqq 1, \qquad |y| \leqq 1 \qquad (i = 1, 2, \cdots, m),$$

since this can always be effected by a substitution of the form

$$x_i = \rho_i x_i', \qquad y = \sigma y' \qquad (i = 1, 2, \cdots, m).$$

Suppose then that K is a number greater than the absolute value of any term in the series $f(1, 1, \cdots, 1, 1)$, that is, greater

than the absolute value of any coefficient in f. If A_0 is the absolute value of a_0, the series

$$F = A_0 y^n - \frac{K y^{n+1}}{1-y} - \frac{K}{1-y} X,$$

where

$$X = \frac{1}{(1-x_1)(1-x_2)\cdots(1-x_m)} - 1,$$

dominates f in the sense that every coefficient except the first has a numerical value equal to or greater than K; and the series B satisfying the relation

$$BF = A_0 y^n + A_1 y^{n-1} + \cdots + A_n$$

analogous to (7) has coefficients numerically greater than the absolute values of those of b. Hence if B converges the same will be true of b.

But it is easy to show that the series B converges. It will certainly do so if convergent series A_k, C, D can be found satisfying the relation

$$A_0 y^n(1-y) - K y^{n+1} - KX = (A_0 y^n + A_1 y^{n-1} + \cdots + A_n)(Cy + D),$$

because then B would have the value

$$B = \frac{1-y}{Cy+D}.$$

On comparing the coefficients of the two highest terms in y in the next to last equation, and for convenience denoting by α the constant value

$$\alpha = -\frac{A_0 + K}{A_0^2},$$

it is found that

$$C = \alpha A_0, \qquad D = 1 - \alpha A_1.$$

By comparing the other powers of y and substituting these values,

we have

$$
\begin{aligned}
A_1 + \alpha A_0 A_2 &= \alpha A_1^2, \\
A_2 + \alpha A_0 A_3 &= \alpha A_1 A_2 \\
\cdots\cdots \\
A_{n-1} + \alpha A_0 A_n &= \alpha A_1 A_{n-1}, \\
A_n &= \alpha A_1 A_n - KX.
\end{aligned}
$$

But these equations have linear terms in $A_1, A_2, \cdots, A_n$ with functional determinant different from zero, and hence have solutions, by the theorems of § 2, which are convergent series in $x_1, x_2, \cdots, x_m$ and have no constant terms.

It is evident, in any numerical case for which f is convergent, that a neighborhood of the origin may be chosen in which the series b is everywhere different from zero. In such a neighborhood all of the values $(x_1, x_2, \cdots, x_m, y)$ which make f vanish are roots of the equation $p = 0$, and vice versa.

If $f(x_1, 0, \cdots, 0, y)$ has its terms of lowest degree homogeneous and of degree n, then the polynomial $p(x_1, 0, \cdots, 0, y)$ has the same initial terms, since the first coefficient of the factor series b is unity.

§ 10. The Zeros of $\varphi(u, v)$, $\psi(u, v)$, or their Functional Determinant

Consider a function $\varphi(u, v)$ whose values in the neighborhood of the origin in the uv-plane are given by a convergent series in u and v which vanishes for $u = v = 0$. If the series contains a factor u in every term it may be written in the form

$$\varphi(u, v) = au^k\Phi(u, v), \tag{9}$$

where a is a constant different from zero and $\Phi(u, v)$ is a convergent series for which $\Phi(0, v)$ has a first term of the form v^m with coefficient unity. According to the results of the preceding section, all of the roots of $\Phi(u, v)$ in a neighborhood of the origin will be roots of a certain polynomial

$$P = v^m + a_1 v^{m-1} + \cdots + a_m, \tag{10}$$

where the coefficients a_k are series in u having no constant terms.

The polynomial P may be equal to the product of two polynomials of similar form,

$$b_0 v^k + b_1 v^{k-1} + \cdots + b_k,$$
$$c_0 v^{m-k} + c_1 v^{m-k-1} + \cdots + c_{m-k},$$

where the coefficients b and c are convergent series in u. In that case the product $b_0 c_0$ must be identically unity, and by dividing the first polynomial by b_0 and multiplying the second by the same series, the two factors will have the form

$$v^k + b_1' v^{k-1} + \cdots + b_k',$$
$$v^{m-k} + c_1' v^{m-k-1} + \cdots + c_{m-k}',$$

The coefficients b' and c' are now series in u without constant terms. Otherwise the product P would have a term of lower degree than v^m, with a coefficient series whose constant term would be different from zero.

It is readily seen from this that the polynomial P is either irreducible in the sense that it can not be decomposed into a product of polynomials of the same sort, or else it is the product of a number of irreducible polynomials of lower degree.

Suppose that $Q(u, v)$ is a polynomial of the form (10) which is irreducible in the sense just described. Then its discriminant with respect to v is a series in u which does not vanish identically, since otherwise Q and Q_v would necessarily have a common factor of the form (10), and Q would not be irreducible. There is a neighborhood $0 < u \leqq u_1$ in which the discriminant is everywhere different from zero, and for any value u satisfying these inequalities the values of v making $Q = 0$ are all distinct. According to the results which have been stated above in the introduction to this chapter of the lectures, the values of v which make Q vanish for different values of u will be defined by m series of the form

(11) $$v = \alpha u^{\mu/\nu} + \alpha' u^{\mu'/\nu} + \cdots;$$

and these series must all be distinct, since for sufficiently small values $u \neq 0$, as has been seen, the roots of Q are all distinct.*

It is evident then that all the roots of $\varphi(u, v)$ in the neighborhood of the origin, including those which correspond to the factor u^k in equation (9), are given by a finite number of elements of the form

$$u = at^p, \qquad v = bt^\mu + b't^{\mu'} + \cdots,$$

where a and b do not vanish simultaneously, and $p, \mu, \mu', \cdots$ are positive integers having no common factor.

The product of factors of the form

$$\{v - \alpha u^{\mu/p} - \alpha' u^{\mu'/p} - \cdots\}, \tag{12}$$

corresponding to the elements of a cycle, is a polynomial $Q_1(u, v)$ of the form (10). For the product Q_1 is a series in $u^{1/p}$ and v which is unchanged when $u^{1/p}$ is replaced by $\omega^\nu u^{1/p}$, and Q_1 must therefore contain only powers of $u^{1/p}$ whose exponents are multiples of p, that is, positive integral powers of u.

On the other hand an irreducible polynomial Q possesses only a single cycle of elements of the form (12). Each element of a cycle belonging to Q gives rise, in fact, to a factor Q_1 of Q of the form (10). The number of elements in the cycle could not be greater than the degree of Q, and neither could it be less, since according to the argument of the paragraph just preceding, Q would then be divisible by a factor of the same form corresponding to the product of the factors (12) belonging to the cycle.

By combining these two results, it follows that *the product of the factors of the form* (12) *corresponding to the elements of a single cycle is an irreducible polynomial of the form* (10), *and conversely the elements of an irreducible polynomial of the form* (10) *form a single cycle.*

The Weierstrassian polynomial P of any function φ is a product of irreducible factors of the same form, some perhaps repeated,

* The method of proof for this statement in the case of a polynomial P is precisely that of the theory of algebraic functions. See the reference above (page 44) to Appell and Goursat.

to each of which there corresponds a cycle of elements. By the order of an element of φ is meant the number of times its factor (12) is repeated in the product $u^k P$. The order is evidently equal to the multiplicity in $u^k P$ of the irreducible factor to which the element belongs. If φ possesses one element of a cycle it must possess the whole cycle. For the polynomial P belonging to φ has then a common factor with the irreducible polynomial Q of the cycle, and so must be divisible by Q.

Suppose now that $\varphi(u, v)$ and $\psi(u, v)$ are two functions of the form described above, and that the functional determinant

$$D(u, v) = \begin{vmatrix} \varphi_u & \varphi_v \\ \psi_u & \psi_v \end{vmatrix} \tag{13}$$

does not vanish identically.

If φ and ψ have an element in common, then they have in common the irreducible polynomial Q of the form (10) *to which the element belongs, and Q is also factor of D.*

The first part of this statement follows from the preceding paragraphs, so that φ and ψ may be supposed to have the forms

$$\varphi = QA, \qquad \psi = QB.$$

When these expressions are substituted in the functional determinant (13) the presence of the factor Q is at once evident.

A similar argument shows that *if φ has an element with corresponding factor Q of multiplicity k, and ψ has the same element and factor with multiplicity l, then D contains the element and its factor with multiplicity $k + l - 1$ at least.*

There is a sort of converse to these statements to the effect that *when φ and D have an element and its factor Q in common, then the element and Q are either multiple in φ or else are common to φ and ψ.*

To prove this let

$$\varphi = QA, \qquad D = QC,$$

and suppose Q not a multiple factor of φ. Then

$$\begin{vmatrix} Q_u A + QA_u & Q_v A + QA_v \\ \psi_u & \psi_v \end{vmatrix} = QC;$$

and it follows readily that the determinant

$$(14) \qquad \begin{vmatrix} Q_u & Q_v \\ \psi_u & \psi_v \end{vmatrix}$$

has the factor Q, since A can not have any element in common with Q. Otherwise it would contain the whole irreducible factor Q.

Since Q is irreducible, its discriminant, a series in u, can not vanish identically, and there is an interval $0 < u \leqq u_1$ in which it is different from zero. For any value of u satisfying these inequalities the polynomials Q and Q_v have no common root. If

$$(15) \qquad u = at^p, \qquad v = \alpha t^\mu + \alpha' t^{\mu'} + \cdots$$

is the parametric form of one of the elements of Q, then $Q(u, v)$ vanishes identically in t when these expressions are substituted, and $Q_v(u, v)$ is not identically zero in t along the element. Hence there is an interval $0 < t < t_1$ in which Q_v is different from zero. Since the determinant (14) has the factor Q and therefore vanishes identically along the curve (15), it follows that

$$Q_v \left(\psi_u \frac{du}{dt} + \psi_v \frac{dv}{dt} \right) = \psi_v \left(Q_u \frac{du}{dt} + Q_v \frac{dv}{dt} \right) = 0$$

is an identity in t. Evidently $\psi(u, v)$ must be constant along the element, and its value is everywhere zero since it vanishes for $t = 0$. Hence ψ has the element (15) in common with Q, and must have Q itself as a factor since Q is irreducible.

The real points (u, v) where one or another of the functions φ, ψ, D vanishes play an important rôle in the investigation which follows. In the discussion of them which follows it will always be understood that when u is real and positive the symbol $u^{1/p}$ stands for the real and positive pth root of u.

If the function φ has no factor u, and if each of its elements when written in the form

$$v = u^{\mu/p}\{\alpha + \alpha' u^{(\mu'-\mu)/p} + \cdots\} \tag{16}$$

has at least one imaginary coefficient, then in a neighborhood of the origin no real point (u, v) with $u > 0$ satisfies the equation $\varphi(u, v) = 0$.

To show this, suppose for the moment that α is imaginary. Then for sufficiently small positive values of u the absolute value of $\alpha' u^{(\mu'-\mu)/p} + \cdots$ will be less than the absolute value of the imaginary part of α, and the parenthesis in the expression (16) will also be imaginary. A similar argument would show v to be complex if one of the higher coefficients were the first not real.

On the other hand, if the coefficients in the expression are all real, then for positive values of u the values of v are real, and the points (u, v) so defined lie on a real arc of the form

$$u = t^p, \qquad v = \alpha t^\mu + \alpha' t^{\mu'} + \cdots \qquad (0 < t < t_1).$$

If the elements of φ are written in the form

$$v = \alpha\epsilon^\mu(-u)^{\mu/p} + \alpha'\epsilon^{\mu'}(-u)^{\mu'/p} + \cdots, \tag{17}$$

where ϵ is a fixed pth root of -1, then an argument similar to that just given shows that $\varphi = 0$ is satisfied by no real points in the neighborhood of the origin with negative values of u, unless at least one of the expressions (17) in $(-u)^{1/p}$ has all of its coefficients real. On the other hand any such element with real coefficients defines points (u, v) on a real arc

$$u = -t^p, \qquad v = \beta t^\mu + \beta' t'^\mu + \cdots \qquad (0 < t < t_1).$$

By combining these results it follows that *all of the real points, in a neighborhood of the origin, which satisfy $\varphi(u, v) = 0$, are the points of a finite number of distinct elements of the form*

$$u = at^p, \qquad v = bt^\mu + b't'^\mu + \cdots \qquad (0 \leqq t < t_1) \tag{18}$$

whose coefficients are real and such that a and b are not both zero.

It may be of interest to note in passing that if an element of φ of the form (16) has real coefficients, then the irreducible polynomial Q which belongs to that element is real. For Q is the product of

$$\{v - \alpha u^{\mu/p} - \alpha' u^{\mu'/p} - \cdots\}$$

and the other factors which arise from it by replacing $u^{1/p}$ by $\omega^\nu u^{1/p} (\nu = 0, 1, 2, \cdots, p-1)$. The coefficients of the product are therefore rational integral functions with real coefficients in the α's and the pth roots of unity, and symmetric in the latter. But symmetric functions of the pth roots of unity are real. A similar remark holds true for the real elements of the form (17).

Two real elements of the form (18) *are said to be distinct if there is an interval* $0 < t \leq t_1$ *on which the points* (u, v) *which they define are all distinct. Any two elements are either distinct or else coincident throughout.*

Let the two elements have the equations

$$u = at^p, \qquad v = bt^\mu + b't^{\mu'} + \cdots \qquad (0 \leq t \leq t_1),$$
$$u = ct^q, \qquad v = dt^\nu + d't^{\nu'} + \cdots \qquad (0 \leq t \leq t_2).$$

If $a = c = 0$ then the elements are distinct unless b and d have the same sign, in which case each defines the same half ray from the origin along the v-axis. If $a = 0$, $c \neq 0$ the elements are distinct. If a and c are both different from zero then the elements are distinct unless the expressions

$$v = b\left(\frac{u}{a}\right)^{\mu/p} + b'\left(\frac{u}{a}\right)^{\mu'/p} + \cdots,$$
$$v = d\left(\frac{u}{c}\right)^{\nu/p} + d'\left(\frac{u}{c}\right)^{\nu'/p} + \cdots,$$

are identical in fractional powers of u, in which case the two elements coincide.

It can readily be seen that if two functions φ and ψ have a real element in common then they must each contain the irreducible real factor which belongs to the element.

§ 11. Singular Points of a Real Transformation of Two Variables

In this section it is proposed to study the singular points of a transformation

$$(19) \qquad x = \varphi(u, v), \qquad y = \psi(u, v)$$

for which φ and ψ are convergent series in u, v with real coefficients. It is presupposed that the functional determinant D of φ and ψ does not vanish identically, and that the real elements of φ and ψ described in § 10 are all distinct. There is an interval $0 \leqq t \leqq t_1$ for which the elements of φ, ψ, and D which are distinct have only the point $(u, v) = (0, 0)$ in common. Some of these elements may belong to both φ and D, or to ψ and D, but none are common to φ and ψ. By further restricting the interval if necessary, it can be effected that the radius

$$\rho = \sqrt{u^2 + v^2}$$

constantly increases on each element as t increases from 0 to t_1. For ρ is a series in t which does not vanish identically, and its derivative has the same character. An interval $0 < t < t_1$ can therefore always be selected on which both ρ and $d\rho/dt$ remain greater than zero.

It follows immediately that a constant ρ_1 can be selected so that any circle about the origin of radius ρ_1 or less is intersected once and but once by each of the elements in question. The real elements of φ, ψ, and D may therefore be represented as shown in Fig. 4.

If the value of ρ_1 is properly restricted then any one of the regions S shown in the figure is transformed in a one-to-one way by the equations (19) *into a region Σ adjoining the origin and lying entirely in one quadrant of the xy-plane. The single-valued inverse functions*

$$(20) \qquad u = f(x, y), \qquad v = g(x, y)$$

so defined are continuous over all of Σ and analytic in its interior.

To prove this consider the functions $r(u, v)$ and $\omega(u, v)$ defined by the equations

$$r = \sqrt{\varphi^2 + \psi^2}, \quad \cos\omega = \frac{\varphi}{r}, \quad \sin\omega = \frac{\psi}{r}.$$

If the radius ρ_1 is properly restricted, then r and ω (modulus 2π) are well defined at every point of the circle with the exception of the origin, since φ and ψ have no real roots in common aside from $(u, v) = (0, 0)$.

The value of r increases monotonically along any analytic curve

$$u = a_1 t + a_2 t^2 + \cdots, \qquad v = b_1 t + b_2 t^2 + \cdots,$$

for which u and v are not identically zero, as may be seen by reasoning similar to that applied above for ρ, after noting that the series for φ and ψ can not vanish identically in t. In particular

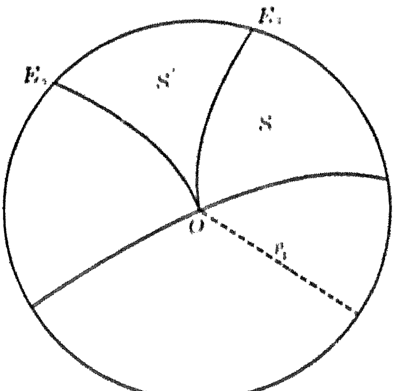

Fig. 4.

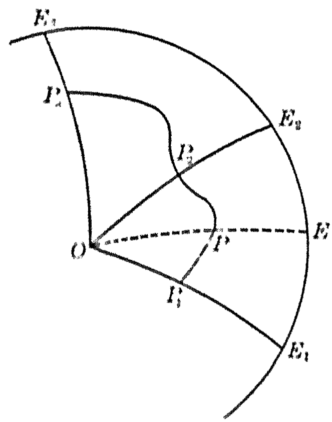

Fig. 5.

if ρ_1 is sufficiently small, then r has this property along the boundaries OE_1 and OE_2 of S, and along an auxiliary arc OE chosen arbitrarily for purposes of proof between the two elements OE_1 and OE_2.

Suppose now that k_1 is the minimum of r along the arc E_1E_2, and select arbitrarily a value k between 0 and k_1. The first of

the equations

$$(21) \qquad r(u, v) = k, \qquad \omega(u, v) = z$$

is satisfied at a unique point $P(u_0, v_0)$ on the arc OE, and the corresponding value of z may be denoted by z_0. The functional determinant of r and ω has the value

$$\frac{\partial(r, \omega)}{\partial(u, v)} = \frac{D(u, v)}{r}$$

and does not vanish anywhere in the interior of S.

The domain in which the equations (21) are to be studied is that consisting of points (u, v, z) for which (u, v) is in S, and z has any real value. According to the first theorem of § 5 and the results of § 2 the equations (21) define two analytic functions

$$(22) \qquad u = u(z), \qquad v = v(z)$$

which take the initial values u_0, v_0 when $z = z_0$, and which may be continued over an interval $z_0 \leqq z < \zeta''$, as described in § 5. If ζ'' is the value defining the largest such interval, the points $(u(z), v(z))$ corresponding to interior points of the interval will all be interior to S, while as z approaches ζ'' the only limit points of the values $(u(z), v(z))$ must lie on the boundary of S. Otherwise the curve (22) could be continued beyond the value ζ''.

The length of the interval $z_0 \leqq z < \zeta''$ is certainly less than $\pi/2$, since in the region S neither $\sin \omega$ nor $\cos \omega$ can vanish. The curve (22) can not intersect itself, since the same values of (u, v) must define the same z by means of the second of equations (21).

As z approaches ζ'', the point $(u(z), v(z))$ approaches a unique limiting point on OE_1 or OE_2. This follows because at any limit point the value of $r(u, v)$ would have to be k, and this can happen at one point P_1 only of PE_2, and at one point P_2 only of OE_2. The curve could not have both P_1 and P_2 as limit points as z approaches ζ'', since then it would necessarily cross the arc OE at the only point P where $r(u, v) = k$, and so would intersect itself.

A similar argument shows that the equations (21) define an arc without double point over an interval $\zeta' < z \leqq z_0$, joining P with that one of the points P_1, P_2 which was not the end of the first arc. For convenience it may be assumed that ζ' is the value belonging to P_1, and ζ'' that for P_2. The preceding inequalities for z would only be reversed if the opposite were the case.

There are no other points in the region S at which $r(u, v) = k$ besides those of the arc P_1P_2 which has just been defined. If there were one not on P_1P_2, it would give rise to a second curve of the same sort joining P_1P_2. But this new curve would necessarily intersect the arc OE at P, and hence must coincide with the original arc P_1P_2 throughout.

For any value $k' < k$ there is a curve similar to P_1P_2 on which all of the points (u, v) making $r(u, v) = k'$ lie.

By means of these results it can now be shown that any two distinct points of the region OP_1P_2 are transformed into two distinct points of the xy-plane. For if (u', v') and (u'', v'') defined the same point (x', y') they would both give $r = \sqrt{x^2+y^2}$ the same value k', and hence must lie on the same curve P_1P_2. But in that case the values of ω corresponding to the two points would necessarily be different, as has been seen above, and hence (x', y') and (x'', y'') could not be the same.

From the final theorem of § 8 it follows at once that the theorem last stated above is true, provided that the circle of radius ρ_1 is altered so that the arc of it which lies between the branches OE_1 and OE_2 lies also within the region OP_1P_2. The region into which S is transformed must lie entirely in one quadrant of the xy-plane, since the values of ω which correspond to points of S are all in one quadrant. In the interior of the image of S the inverse functions (20) are analytic, since at interior points of S the determinant D is different from zero.

Some conclusions with regard to the distribution of the elements of φ, ψ, and D can be readily derived from the discussion just preceding. For example, no region S can be bounded

by two elements of φ. If it were not so, then in a region bounded by two elements of φ the value of ω on the branch OE_1 would be everywhere $\pi/2$, or else everywhere $-\pi/2$, and the same is true for OE_2. But this is impossible since along the arc P_1P_2 the value of ω varies monotonically through an interval less than $\pi/2$. A similar remark holds for the elements of ψ. Hence it follows easily that

Between any elements of D the elements of φ and ψ, if there are any, must separate each other.

If the determinant D has opposite signs in two adjoining regions S and S' of the circle of radius ρ_1 in the uv-plane, shown in Fig. 5, their transforms in the xy-plane will be folded over the image of the curve OE_2 and will overlap. In order to prove this, let it first be remembered that along the element OE_2

$$\frac{dr}{dt} = r_u\frac{du}{dt} + r_v\frac{dv}{dt} \neq 0,$$

so that r_u and r_v can not vanish at any point P_2 different from the origin. Neither can they vanish at an interior point of one of the regions S, since at a point where

$$r_u = \frac{\varphi\varphi_u + \psi\psi_u}{r} = 0, \qquad r_v = \frac{\varphi\varphi_v + \psi\psi_v}{r} = 0,$$

the determinant D would necessarily have the value zero, and this does not occur in the interior of S. The equations

$$r_u\frac{du}{dz} + r_v\frac{dv}{dz} = 0, \qquad \omega_u\frac{du}{dz} + \omega_v\frac{dv}{dz} = 1$$

are satisfied everywhere between P_1 and P_2 on the arc (22). Hence

$$\frac{du}{dz} = -\frac{r}{D}r_v, \qquad \frac{dv}{dz} = \frac{r}{D}r_u.$$

As z approaches ζ'' the direction cosines of the tangent to the

so determined are continuous at all points of the sheet Σ and analytic in the interior of Σ. If in two adjoining regions, say S_1 and S_2, the signs of D are opposite, then the images Σ_1 and Σ_2 overlap in the neighborhood of their common boundary $O\pi_2$; if the signs of D are the same, the regions Σ_1 and Σ_2 adjoin along $O\pi_2$ without overlapping.

The adjoining figure illustrates the case when D has four real elements and the signs of D are opposite in any two adjoining regions S. Further illustrations of the theorem are given in § 14.

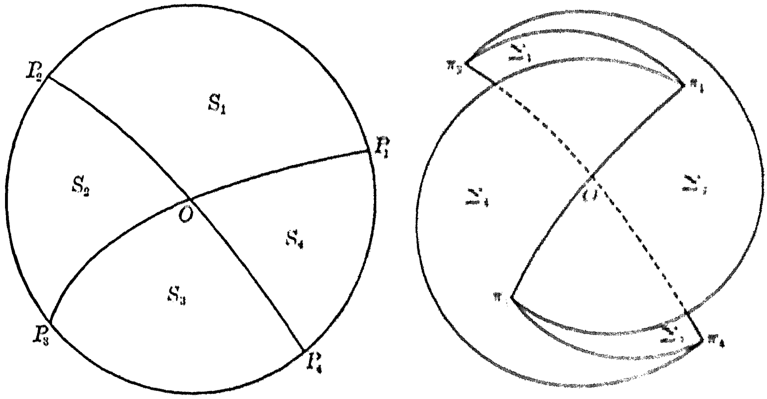

FIG. 6.

It has not been proved above that the functions (24) are continuous on a boundary $O\pi$ of one of the regions Σ. Suppose that π is a point of such a boundary, and let

$$(25)\qquad \pi_1,\ \pi_2,\ \pi_3,\ \cdots$$

be any sequence of points of Σ with limit π. The corresponding points

$$(26)\qquad p_1,\ p_2,\ p_3,\ \cdots$$

of S have condensation points in S, one of which may be denoted by p. There is then a sub-sequence

$$p_1',\ p_2',\ p_3',\ \cdots$$

among the points (26) whose limit is p; and on account of the continuity of the functions (23), the corresponding points

$$\pi_1', \; \pi_2', \; \pi_3', \; \cdots \tag{27}$$

of the sequence (25) must have as limit point the image of p in Σ. But the limit of (27) is necessarily π, and π is therefore the image of p. It follows at once that the sequence (26) has a unique limit point p which is the image of π, and from this property the continuity of the functions (24) in the ordinary sense can be readily deduced.

The functions φ, ψ, and D can be expanded in the form

$$\begin{aligned} \varphi &= \varphi_m + \varphi_{m+1} + \cdots, \\ \psi &= \psi_n + \psi_{n+1} + \cdots, \\ D &= D_{m+n-2} + D_{m+n-1} + \cdots, \end{aligned} \tag{28}$$

where φ_k, ψ_k, D_k are homogeneous polynomials in u, v of degree k, and

$$D_{m+n-2} = \begin{vmatrix} \dfrac{\partial \varphi_m}{\partial u} & \dfrac{\partial \varphi_m}{\partial v} \\ \dfrac{\partial \psi_n}{\partial u} & \dfrac{\partial \psi_n}{\partial v} \end{vmatrix}.$$

If the real roots of φ_m, ψ_n, and D_{m+n-2} are all simple roots and distinct from each other, there will be an element of φ, ψ, or D in each of the corresponding directions, and a notion of the character of the transformation can be derived without difficulty. In the applications of § 14 this remark is of frequent service.

§ 12. The Case where the Functional Determinant Vanishes Identically

It is well known that when the functional determinant of two analytic functions φ and ψ vanishes identically, then near any point where not all of the derivatives φ_u, φ_v, ψ_u, ψ_v vanish the functions φ and ψ satisfy a relation of the form

$$F(\varphi, \psi) = 0$$

identically in u and v. It is possible to show that such a relation exists also near a singular point at which the four derivatives above all vanish.

If a relation can be found after a substitution of the form

$$(29) \qquad u = \alpha u_1 + \beta v_1, \qquad v = \gamma u_1 + \delta v_1,$$

for which $\alpha\delta - \beta\gamma$ does not vanish, then it will surely be satisfied when u_1 and v_1 are replaced by the original variables u, v.

Suppose then that the analytic functions φ and ψ have already been prepared by a transformation (29) in such a way that in the expansions (28) φ_m and ψ_n both contain terms in u alone. By applying the preparation theorem of Weierstrass to the functions $\varphi(u, v) - x$ and $\psi(u, v) - y$, two polynomials

$$P(u, v, x) = u^m + a_1 u^{m-1} + \cdots + a_m,$$
$$Q(u, v, y) = u^n + b_1 u^{n-1} + \cdots + b_n$$

are obtained, whose coefficients are convergent series, without constant terms, in v, x and v, y, respectively. In a certain vicinity

$$|x| < \epsilon, \quad |y| < \epsilon, \quad |u| < \epsilon, \quad |v| < \epsilon$$

the only solutions of the equations

$$(30) \qquad \varphi(u, v) - x = 0, \qquad \psi(u, v) - y = 0$$

are values (u, v, x, y) which make P and Q vanish also, and vice versa.

The resultant of P and Q is a convergent series $R(v, x, y)$ for which $R(0, x, y)$ does not vanish identically. For if all of the coefficients of $R(0, x, y)$ were zero, there would be a region

$$(31) \qquad v = 0, \quad |x| < \delta, \quad |y| < \delta \qquad (\delta < \epsilon)$$

at any point of which the polynomials P and Q have a common root in absolute value less than ϵ, and the set of values $(u, 0, x, y)$ so defined satisfies also the equations (30). The existence of such a region is, however, impossible, since when y' is given

satisfying (31), a value x' can always be selected which is different from the values of $\varphi(u, 0)$ at all of the n roots of $Q(u, 0, y')$. For such a set $v = 0, x', y'$ in the region (31) there would be no corresponding value u' satisfying the equations (30).

The resultant $R(v, x, y)$ vanishes identically in u, v when x and y are replaced by φ and ψ. For R is expressible in the form

$$R(v, x, y) = MP + NQ,$$

where M and N are polynomials in u with coefficients which are series in v, x, y, and P and Q vanish identically when $x = \varphi$, $y = \psi$.

The series $R(0, \varphi, \psi)$ vanishes identically in u, v. If not, there would be a straight line $u = kv$ on which $R(0, \varphi, \psi)$ and φ_u are different from zero except at the origin. Let (u', v') be a point of this line near $(u, v) = (0, 0)$, at which φ and ψ have the values φ' and ψ', respectively. The series

$$R(0, \varphi, \psi) + R_v(0, \varphi, \psi)v + \cdots \tag{32}$$

vanishes identically, in particular along the curve

$$\varphi(u, v) = \varphi' \tag{33}$$

through the point (u', v'). Since φ_u does not vanish at (u', v'), this curve can be expressed in the form

$$u = U(v),$$

and along it

$$\frac{d}{dv}\psi(U, v) = \left[\psi_u \frac{-\varphi_v}{\varphi_u} + \psi_v\right]_{u = U(v)} = 0,$$

since the functional determinant of φ and ψ vanishes identically. On the curve (33) the function ψ has therefore the constant value ψ', and the series (32) takes the form

$$R(0, \varphi', \psi') + R_v(0, \varphi', \psi')v + \cdots$$

and vanishes identically in v. Its coefficients must therefore all vanish, since a series whose zeros have a point of condensation

in the interior of its circle of convergence must have all of its coefficients equal to zero. This contradicts, however, the assumption that a point (u', v') exists at which $R(0, \varphi, \psi)$ does not vanish.

It has been shown therefore that *in case the functional determinant of the two convergent series*

$$\varphi = \varphi_m + \varphi_{m+1} + \cdots,$$
$$\psi = \psi_n + \psi_{n+1} + \cdots$$

vanishes identically, the two functions φ, ψ *satisfy a relation of the form*

$$F(\varphi, \psi) = 0$$

identically in u, v, *where* F *is itself a convergent series in its two arguments. This statement is true even when* φ *and* ψ *both have singular points at the origin.*

It is evident that when $D = 0$ the transformation

$$x = \varphi(u, v), \qquad y = \psi(u, v)$$

makes all of the points in the neighborhood of the origin in the uv-plane correspond to points on the various branches of the curve

$$F(x, y) = 0$$

in the xy-plane. The points (x, y) which are obtained by the transformation do not cover any region.

§ 13. A Generalization of the Preparation Theorem of Weierstrass

Consider for a moment two functions

$$(34) \qquad f(u, v, x_1, x_2, \cdots, x_m), \quad g(u, v, x_1, x_2, \cdots, x_m)$$

which are polynomials in the variables u, v and have for coefficients convergent series in $x_1, x_2, \cdots, x_m$. According to the usual algebraic theory of elimination, there exists a polynomial p in v

which has convergent series in the x's as coefficients, and which is linearly expressible in the form

$$p = cf + dg,$$

where c and d are polynomials of the same character as f and g. If a set of variables (u, v, x) make f and g both vanish, then v must be a root of the polynomial p; and conversely to any root of p corresponding to given values x, there exists at least one pair of values (u, v) which satisfy the two equations $f = g = 0$.

There is a generalization of the preparation theorem of Weierstrass from which similar results may be deduced with respect to two functions f and g which are not polynomials but series in the variables u and v, and with respect to the roots of such functions in a neighborhood of any set of values (u_0, v_0, x_0) making f and g vanish. As in the proof of the theorem of § 9, the point in whose neighborhood f and g are to be studied may be taken without loss of generality at the origin.

Suppose then that f and g are two convergent series in u, v, x vanishing for $(u, v, x) = (0, 0, 0)$, and such that $f(u, v, 0, 0, \cdots, 0)$ and $g(u, v, 0, 0, \cdots, 0)$ have no common factor. Then there exists a polynomial

$$p = v^n + p_1 v^{n-1} + \cdots + p_n, \tag{35}$$

in which the coefficients p_k $(k = 1, 2, \cdots, n)$ are convergent series in x having no constant terms, with the following properties: (1) *it is linearly expressible in the form*

$$p = cf + dg,$$

where c and d are convergent power series in u, v, x; (2) *in a properly chosen neighborhood*

$$|u| < \epsilon, \quad |v| < \epsilon, \quad |x| < \epsilon \tag{36}$$

every root (u, v, x) of f and g must also make p vanish; (3) *there exists a constant $\delta \leqq \epsilon$ such that for any x in the region*

$$|x| < \delta \tag{37}$$

there is associated with each root v of p a solution (u, v, x) of the equations $f = g = 0$ satisfying the inequalities (36).*

If $f(u, v, 0, 0, \cdots, 0)$ and $g(u, v, 0, 0, \cdots, 0)$ have no common factor, then one at least of them, say f, has terms in the variable u alone, and according to the preparation theorem of Weierstrass $f(u, v, x)$ has as factor a polynomial of the form

$$(38) \qquad a_0u^m + a_1u^{m-1} + \cdots + a_{m-1}u + a_m = hf,$$

in which a_0 is a constant different from zero, and $a_1, a_2, \cdots, a_m$ are series in v, x without constant terms. The symmetric functions of the roots $u_1, u_2, \cdots, u_m$ of this polynomial are expressible rationally and integrally in terms of the coefficients $a_1, a_2, \cdots, a_m$, and are therefore convergent series in v, x. The product

$$(39) \qquad \prod_{k=1}^{m} g(u_k, v, x) = h(v, x)$$

is a convergent series in u_k, v, x, also symmetric in the variables u_k, and hence expressible as convergent series in v, x.

The function $h(v, 0)$ does not vanish identically, on account of the hypothesis that $f(u, v, 0, 0, \cdots, 0)$ and $g(u, v, 0, 0, \cdots, 0)$ have no common factor. If it did vanish identically, then for every sufficiently small value of v one at least of the expressions $g(u_k, v, 0)$ would vanish. But in § 10 it was seen that when $f(u, v, 0)$ and $g(u, v, 0)$ have no factor in common, there is always an interval $0 < v \leqq v_1$ in which there is no value v belonging to a pair (u, v) making both of these functions vanish.

The preparation theorem of Weierstrass can therefore be applied also to the function $h(v, x)$, and the polynomial so found is the one desired in the theorem. For, in the first place, a constant ϵ can be chosen so small that every root (u, v, x) of f and g in the region (36) must be one of the sets (u_k, v, x), and must make

* A proof that the values of u and v belonging to the roots of a system of equations of the form (34) are roots of polynomials similar to (35) was given by Poincaré in the introduction to his Thesis, "Sur les propriétés des fonctions définies par les équations aux différences partielles," Paris (1879).

the product (39), and hence p, vanish. In the second place, a constant $\delta \leqq \epsilon$ can be taken so small that every root v of p as well as the corresponding sets (u_k, v, x) lie in the domain (36). One at least of these sets must evidently satisfy $g = 0$ as well as $f = 0$. The restrictions on δ and ϵ have been stated somewhat roughly, but the reader will readily convince himself that these quantities may be selected so that the convergence of the different series and their equivalence with the corresponding polynomials are properly adjusted.

Finally, the polynomial p is linearly expressible in the form described in the theorem, in terms of f and g. To prove this, suppose that the above process has been applied to the functions $f - \alpha$ and $g - \beta$. A polynomial $P(v, x, \alpha, \beta)$ with coefficients which are series in x, α, β is then found, which may be written in the form

$$P(v, x, \alpha, \beta) = P(v, x, 0, 0) + C\alpha + D\beta,$$

where C and D are convergent series in the arguments of P. The series $P(u, x, f, g)$ vanishes identically in u, v, x since $P = 0$ must be satisfied by every set of variables (u, v, x, α, β) in a neighborhood of the origin which make $f - \alpha$ and $g - \beta$ vanish, certainly then by the set (u, v, x, f, g). Hence

$$P(v, x, 0, 0) = -Cf - Dg$$

is an identity in u, v, x, when α and β are replaced in C and D by the series f, g. But $P(v, x, 0, 0)$ is precisely the polynomial $p(v, x)$ found above, since for $\alpha = \beta = 0$ the steps in the construction of $P(v, x, 0, 0)$ are identical with those used in finding p.

*If the series $f(u, v, 0, 0, \cdots, 0)$ and $g(u, v, 0, 0, \cdots, 0)$ begin with homogeneous polynomials having no common factor of degrees m and n, respectively, then the degree of the polynomial p is $\nu = mn$.**

* In a paper of recent date the writer has developed a generalization of this theorem and the results which follow, for a system of equations of the form $f_i(x_1, x_2, \cdots, x_m; y_1, y_2, \cdots, y_n) = 0$ $(i = 1, 2, \cdots, n)$. See *Transactions of the American Mathematical Society*, vol. 13 (1912), p. 133.

Let the lowest terms of $f(u, v, 0, 0, \cdots, 0)$ and $g(u, v, 0, 0, \cdots, 0)$ be denoted by $\varphi_m(u, v)$ and $\psi_n(u, v)$, respectively. One of the two, say φ_m, has a term involving u alone with coefficient different from zero, since φ_m and ψ_n have no common factor. The terms of lowest degree in the polynomial (35) are also φ_m, since the series b has constant term unity. In the product (39) the terms may be rearranged into groups of the form $cv^\rho U$, where U is a homogeneous symmetric function of a certain degree σ in $u_1, u_2, \cdots, u_m$. The expression for such a symmetric function is isobaric and has the weight σ in the coefficients of the polynomial (35). When $x = 0$ the terms of lowest degree in U will be at least of degree σ in v, since each coefficient a_k of (35) begins with the coefficient of u^{m-k} in the polynomial $\varphi_m(u, v)$. The terms of lowest degree in v alone in the product (39) will therefore be those of the product

$$\prod_{k=1}^{m} \psi_n(u_k, v),$$

and they have the value $v^{mn}R/a_0$, in which a_0 is the coefficient of v^m in $\varphi_m(u, v)$ and R is the resultant of $\varphi_m(1, v)$ and $\psi_n(1, v)$.* But since φ_m and ψ_n have no common factor the coefficient of v^{mn} is surely different from zero, and the theorem last stated follows at once.

If the substitution

$$v = -tu + z$$

is made, in which t is a new variable, the series

$$(40) \qquad \begin{aligned} F(u, z, x, t) &= f(u, z - tu, x), \\ G(u, z, x, t) &= g(u, z - tu, x) \end{aligned}$$

have a polynomial

$$(41) \qquad P(z; x, t) = CF + DG$$

with properties similar to those of p and of the same degree ν. In

* See, for example, König, Einleitung in die allgemeine Theorie der algebraischen Grössen, p. 311 and p. 271 (d).

a properly chosen region

(42) $$|u| < \epsilon, \quad |v| < \epsilon, \quad |x| < \epsilon$$

every root (u, v, x) of f and g defines a factor $z - tu - v$ of P. If $\delta < \epsilon$ is sufficiently small and x a set of variables satisfying

(43) $$|x| < \delta,$$

then P has ν factors of the form $z - tu - v$, for each of which the values (u, v, x) are a solution of the equations $f = g = 0$ in the region (42).

The degree of P must be the same as that of p, since for $x = t = 0$ the series $F(u, z, 0, 0)$, $G(u, z, 0, 0)$ are identically equal to the series $f(u, v, 0)$ and $g(u, v, 0)$ when v is replaced by z. In a certain region

(44) $$|u| < \epsilon_1, \quad |z| < \epsilon_1, \quad |x| < \epsilon_1, \quad |t| < \epsilon_1,$$

where ϵ_1 is for convenience taken less than unity, every root system (u, z, x, t) of F and G makes P vanish also. If ϵ is taken less than $\epsilon_1/2$ and t is restricted to the range $|t| < \epsilon_1$, every root system (u, v, x) of f and g in the region (42) gives values u, $z = tu + v$, x, t satisfying the inequalities (44), and hence P must vanish identically in t and have $z - tu - v$ as a factor.

Suppose then that ϵ is a constant satisfying the requirements of the theorem with respect to the region (42), and that the region analogous to (37) for the polynomial P and the constant $\epsilon/2$ is

(45) $$|x| < \delta, \qquad |t| < \delta;$$

and let $x = \xi$ be any set of values satisfying these inequalities. If the discriminant of P is not identically zero in t for $x = \xi$, a value $t = \tau$ can be selected also satisfying (45) and such that all the roots z of P corresponding to the values ξ, τ are distinct. There are then ν distinct root systems (u, z, ξ, τ) satisfying the inequalities (41) with ϵ_1 replaced by $\epsilon/2$. The corresponding values $(u, v = z - tu, \xi)$ are ν distinct roots of f and g lying in the region (41). According to the paragraph just preceding, P has therefore ν distinct factors $z - tu - v$.

In case the discriminant of P vanishes identically in t for $x = \xi$, the multiple factors of $P(z; \xi, t)$ can be separated out by the highest common divisor process, and the factorization of the resulting polynomial can then be discussed in a manner similar to that just explained. In either case, therefore, $P(z; \xi, t)$ has only linear factors of the form $z - tu - v$.

The number and character of the root systems (u, v, x) of the functions f and g in the neighborhood of the origin are well defined by means of the polynomial $P(z; x, t)$. To any x in the region (43) there correspond ν root systems (u, v, x) not necessarily all distinct, and the ν-valued functions $u(x)$, $v(x)$ so defined are continuous. This is evidently true for the function $v(x)$, since its values are the roots of the polynomial $P(v; x, 0)$ whose coefficients are analytic in x. Similarly z is continuous in x, t, since its values are the roots of $P(z; x, t)$, and it follows that $u = (z - v)/t$, for a fixed value $t \neq 0$, must be continuous in x.

If P is not irreducible, that is, not decomposable into similar factors of lower degrees, its discriminant $\Delta(x, t)$ can not vanish identically in x, t. At any value $x = \xi$ where $\Delta(\xi, t)$ is not identically zero in t, the ν factors $z - tu - v$ of P are all distinct. If $t = \tau$ is selected so that $\Delta(\xi, \tau) \neq 0$, the roots of P are distinct analytic functions of x and t in the neighborhood of ξ, τ, and the corresponding values of u and v are analytic functions of x in the neighborhood of ξ.

The values $x = \xi$ near which the ν-valued functions u, v do not surely have ν distinct analytic branches, are those for which $\Delta(\xi, t)$ vanishes identically in t. At such a point some of the values of the root-systems (u, v) coincide, and only those which are distinct belong necessarily to analytic branches of the functions u, v. The values ξ which make $\Delta(\xi, t)$ identically zero must belong to one of the totalities of points defined by equating to zero the coefficients of the finite number of powers of t in the discriminant $\Delta(x, t)$.*

* For the characterization of these totalities after the method of Kronecker for algebraic equations, see Kistler, "Ueber Funktionen von mehreren komplexen Veränderlichen," Dissertation, Göttingen, 1905.

If $P(z, x, t)$ is reducible, arguments similar to those above can be applied to any one of its irreducible factors.

The multiple roots (u, v, x) of the functions f and g are characterized by the property that the functional determinant $\partial(f, g)/\partial(u, v)$ is zero at such points.

For from the identity (44) in u, z, x, t, it follows by differentiation that

$$(46) \qquad \begin{aligned} 0 &= C_u F + D_u G + C F_u + D G_u, \\ P_z &= C_z F + D_z G + C F_z + D G_z. \end{aligned}$$

If the determinant

$$\begin{vmatrix} F_u & F_z \\ G_u & G_z \end{vmatrix} = \begin{vmatrix} f_u & f_v \\ g_u & g_v \end{vmatrix}$$

vanishes at a solution (u, v, x) of $f = g = 0$, the two equations above show that

$$CF_u + DG_u = 0, \qquad P_z = CF_z + DG_z = 0$$

for the values $(u, z = -tu + v, x)$; and it follows that $z - tu - v$ is a multiple factor of P, since it occurs also in P_z.

On the other hand suppose that at a set of values (u', v', ξ) the determinant $\partial(f, g)/\partial(u, v)$ is different from zero, while f and g vanish. It is to be shown that the polynomial $P(z; \xi, t)$ has $tu' + v'$ as a simple root. All of the roots of $P(z; \xi, t)$ have the form $tu + v$, and some are perhaps multiple. Those which are distinct will remain distinct for a numerical value $t = \tau$ if τ is properly selected, and the derivative

$$(47) \qquad F_u(u', \zeta, \xi, \tau) = f_u(u', v', \xi) - \tau f_v(u', v', \xi)$$

can at the same time be made different from zero, ζ being the expression $\tau u' + v'$. In the expressions

$$(48) \qquad A_0 u^m + A_1 u^{m-1} + \cdots + A_{m-1} u + A_m = BF,$$

$$(49) \qquad \prod_{k=1}^{m} G(u_k, z, \xi, \tau) = H(z, \xi, \tau),$$

analogous to (38) and (39) for the functions $F(u, z, \xi, t)$ and $G(u, z, \xi, t)$, the factor $G(u_1, z, \xi, \tau)$, where u_1 is the root of (48) which reduces to u' for $z = \zeta$, is the only one which vanishes for $z = \zeta$. To prove this it can be seen in the first place that u' is a simple root of (48) for $z = \zeta$, since the derivative (47) is different from zero. Furthermore when $z = \zeta$ no other root u_2 distinct from u_1 can make $G(u_2, z, \xi, \tau)$ vanish. Otherwise f and g would vanish not only at the values (u', v', ξ), but also at $(u_2', \zeta - \tau u_2', \xi)$, where u_2' is the value of u_2 for $z = \zeta$; and $P(z; \xi, t)$ would have two roots, $tu' + v' = tu' + \zeta - \tau u'$ and $tu_2' + \zeta - \tau u_2'$, which are distinct for $t \neq \tau$ and equal to ζ when $t = \tau$. On account of the way in which τ was selected, this is impossible.

The root u_1 of (48), that is to say also of F, has an expansion of the form

$$u_1 - u' = -\frac{F_z(u', \zeta, \xi, \tau)}{F_u(u', \zeta, \xi, \tau)}(z - \zeta) + \cdots$$

in powers of $z - \zeta$; and the value of $G(u_1, z, \xi, \tau)$ is a series

$$\frac{F_u G_z - F_z G_u}{F_u}(z - \zeta) + \cdots$$

whose first term is different from zero, since for the values (u', ζ, ξ, τ) we have

$$\begin{vmatrix} F_u & F_z \\ G_u & G_z \end{vmatrix} = \begin{vmatrix} f_u(u', v', \xi) & f_v(u', v', \xi) \\ g_u(u', v', \xi) & g_v(u', v', \xi) \end{vmatrix} \neq 0,$$

as is readily seen from equations (40). Hence the quotient $H(z, \xi, \tau)/(z - \zeta)$ is different from zero, and neither $H(z, \xi, t)$ nor its polynomial $P(z; \xi, t)$ can have more than one factor $z - tu' - v'$.

§ 14. Applications of the Preceding Theory

The real transformation

$$(50) \qquad \begin{aligned} x &= \varphi(u, v) = a_{10}u + a_{01}v + a_{20}u^2 + \cdots, \\ y &= \psi(u, v) = b_{10}u + b_{01}v + b_{20}u^2 + \cdots, \end{aligned}$$

has a singular point at the origin when

$$(51) \qquad \begin{vmatrix} a_{10} & a_{01} \\ b_{10} & b_{01} \end{vmatrix} = 0.$$

If one of the elements of the determinant is different from zero, it may be assumed without loss of generality to be a_{10}; then after two transformations

$$u' = a_{10}u + a_{01}v, \qquad v' = v,$$
$$x' = x, \qquad y' = -\frac{b_{10}}{a_{10}}x + y$$

the equations (50) take the form

$$(52) \qquad \begin{aligned} x &= u + a_{20}u^2 + a_{11}uv + a_{02}v^2 + \cdots, \\ y &= \phantom{u + {}} b_{20}u^2 + b_{11}uv + b_{02}v^2 + \cdots. \end{aligned}$$

For convenience the primes have been dropped, and the notation for coefficients of terms of higher degree than the first is the same as that in the original equation. It may further be supposed that the polynomials

$$\varphi_1 = u, \qquad \psi_2 = b_{20}u^2 + b_{11}uv + b_{02}v^2$$

have no common factor, in other words that $b_{02} \neq 0$. The origin is then a singular point for the transformation (50) of a very general type, since aside from the assumption (51) only inequalities on the coefficients of the series have been exacted.

The functional determinant has the expansion

$$D(u, v) = b_{11}u + 2b_{02}v + \cdots,$$

and hence has a single branch

$$v = -\frac{b_{11}}{b_{02}}u + \cdots,$$

along which D vanishes and on opposite sides of which D has different signs. The image Δ of this curve in the xy-plane has

an ordinary point at the origin, as shown by its equations

$$x = u + \cdots, \qquad y = \frac{4b_{02}b_{20} - b_{11}^2}{b_{02}} u^2 + \cdots.$$

The region S in the figure has in it one real element of φ and at most two of ψ, since the solutions of $\varphi = 0$ lie on a single real

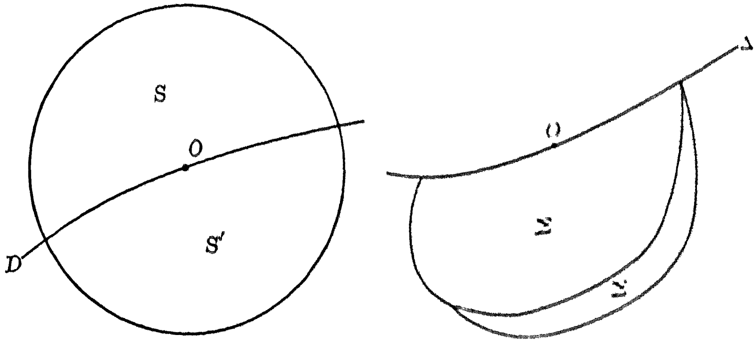

FIG. 7.

curve through the origin, and those of $\psi = 0$ are either imaginary or else lie on two real branches. Hence the region Σ which is the image of S lies on one side only of the curve Δ and overlaps the image Σ' of S'.

Since φ_1 and ψ_2 have no common factor, the theorems of § 13 show that there exist two constants, δ and ϵ, such that the equations (52) have two and only two solutions $[u_1(x, y), v_1(x, y), x, y]$, $[u_2(x, y), v_2(x, y), x, y]$ in the region

$$|u| < \epsilon, \qquad |v| < \epsilon, \qquad |x| < \epsilon, \qquad |y| < \epsilon$$

corresponding to any (x, y) in the region

$$|x| < \delta, \qquad |y| < \delta.$$

The functions u_1, v_1, u_2, v_2 so defined are everywhere continuous and the two solutions above are analytic and distinct except along the curve Δ. On one side of Δ they are imaginary, on the other real.

Another interesting case is that of a transformation (50) for which again the coefficients are real, and

$$\frac{\partial\varphi}{\partial u}=\frac{\partial\psi}{\partial v}, \qquad \frac{\partial\varphi}{\partial v}=-\frac{\partial\psi}{\partial u}.$$

Such a transformation might be called a monogenic transformation. It follows at once that φ and ψ must begin with two homogeneous polynomials, φ_m and ψ_m, of the same degree m, which also satisfy the last equations. Consequently

$$\varphi_m+i\psi_m=(a+ib)(u+iv)^m=\rho^m(a+ib)(\cos\theta+i\sin\theta)^m$$

and

$$\varphi_m=\rho^m(a\cos m\theta-b\sin m\theta), \quad \psi_m=\rho^m(a\sin m\theta+b\cos m\theta),$$

where a and b are not both zero. These equations show that $\varphi_m(u, v)$ and $\psi_m(u, v)$ have each m real linear factors in u, v, and that no factor of φ_m is also in ψ_m.

The determinant $D(u, v)$ has an expansion

$$D(u, v)=D_{2m-1}+D_{2m}+\cdots,$$

where

$$D_{2m-1}=\begin{vmatrix}\dfrac{\partial\varphi_m}{\partial u} & \dfrac{\partial\varphi_m}{\partial v}\\[2ex] \dfrac{\partial\psi_m}{\partial u} & \dfrac{\partial\psi_m}{\partial v}\end{vmatrix}=\left(\frac{\partial\varphi_m}{\partial u}\right)^2+\left(\frac{\partial\varphi_m}{\partial v}\right)^2.$$

The homogeneous polynomial D_{2m-1} has no real root, since such a root would necessarily belong to both $\partial\varphi_m/\partial u$ and $\partial\varphi_m/\partial v$, and from the equations

$$m\varphi_m=u\frac{\partial\varphi_m}{\partial u}+v\frac{\partial\varphi_m}{\partial v}, \quad m\psi_m=u\frac{\partial\psi_m}{\partial u}+v\frac{\partial\psi_m}{\partial v}=-u\frac{\partial\varphi_m}{\partial v}+v\frac{\partial\varphi_m}{\partial u}$$

it follows that φ_m and ψ_m would then have a common factor. Hence there are no real points at which D vanishes near the origin in the uv-plane.

7

The argument of § 11 shows that the elements of φ_m and ψ_m separate each other and that a neighborhood of the origin in the uv-plane is transformed into a sheet winding m times around the origin in the xy-plane, as shown in the figure. This is the

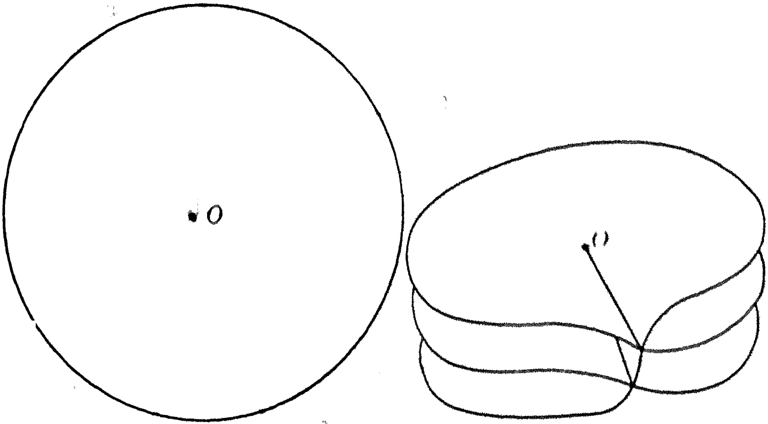

FIG. 8.

well-known transformation of the neighborhood of the origin in a complex w-plane by means of a relation of the form

$$z = Aw^m + A'w^{m+1} + \cdots,$$

where $z = x + iy$ and $w = u + iv$. The figure is drawn for $m = 3$.

There are many other special cases similar to those just given which might be elucidated by means of the theorems of the preceding sections, but for which the methods in the two examples just given are typical. It may be of interest, however, to exhibit an example which illustrates the use of the theorems of § 8, as well as the behavior of a transformation at singular points.

Suppose that the real uv-plane is transformed by means of the equations

$$(53) \qquad x = \frac{u^2}{2} - uv + \frac{v^2}{2} + \frac{u^3}{3}, \qquad y = \frac{u^2}{2} + uv + \frac{v^2}{2}.$$

The functional determinant has the value

$$D(u, v) = (u + v)(u^2 + 2u - 2v)$$

and it vanishes along the curves

$$v = -u, \qquad v = u + \frac{u^2}{2},$$

which have, respectively, the images

$$(54) \qquad \begin{aligned} x &= 2u^2 + \frac{u^3}{3}, & y &= 0, \\ x &= \frac{u^3}{3} + \frac{u^4}{8}, & y &= \frac{u^2}{8}(u+4)^2 \end{aligned}$$

in the xy-plane. These curves are shown in the accompanying

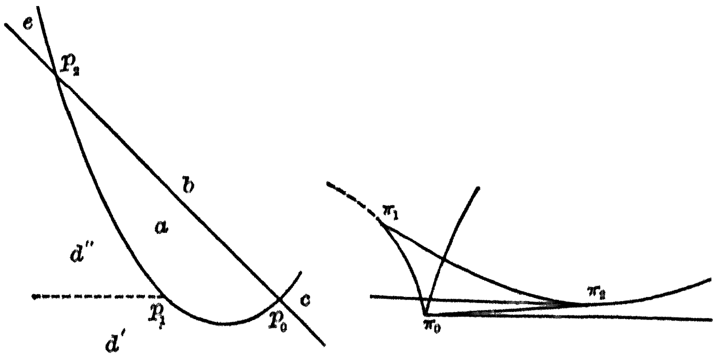

FIG. 9.

figures, the x-axis being drawn triply between $x = 0$ and $x = 32/3$ since this segment is described three times by the point (54) with varying u. To the auxiliary arc $-\infty < u \leqq -2$, $v = 0$ there corresponds the curve

$$x = \frac{u^2}{2} + \frac{u^3}{3}, \qquad y = \frac{u^2}{2} \quad (-\infty < u \leqq -2)$$

shown dotted in the figure.

Consider now, for example, the region a in the uv-plane.

Its boundary is transformed into the boundary of the region α in Fig. 10. According to the generalization of the theorem of

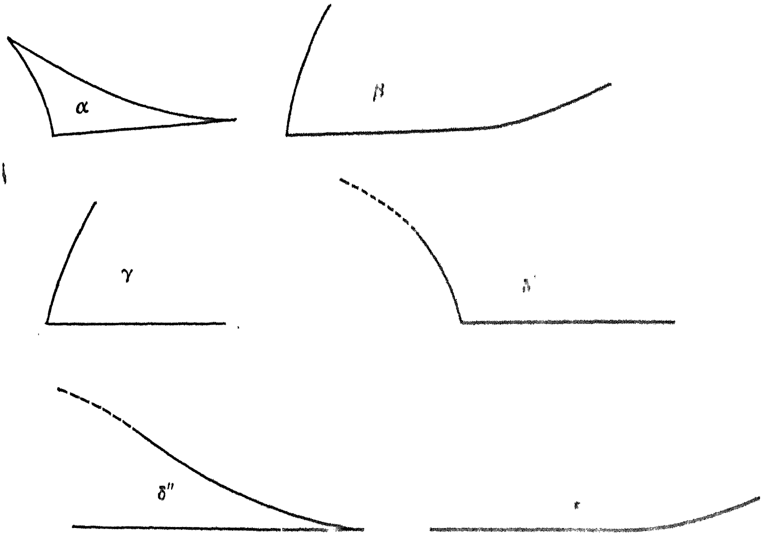

FIG. 10.

Schoenflies in § 8, the transformation defines a one-to-one correspondence between the regions a and α; and the inverse functions $u(x, y)$, $v(x, y)$ so defined are continuous over α and analytic in its interior.

Consider now the region of points (u, v, x, y) defined by the conditions that (u, v) shall lie in the region b or on its boundary, while (x, y) is unrestricted. There is but one sheet of solutions of equations (53) in this region, since any two particular solutions (u', v', x', y'), (u'', v'', x'', y'') interior to the sheet can be joined by a continuous curve lying entirely within the sheet, as may be seen by joining (u', v'), (u'', v'') by a continuous curve in b. No one of the solutions in question has a projection (x, y) outside of β, since otherwise every point exterior to β would be such a projection, according to the third theorem of § 5 or the fourth of § 8; and from the second of equations (53) it is evident that no solution (u, v, x, y) has a negative value for y. On the other

hand every point of β is the projection of a solution. Since β is simply connected, it follows from the fourth theorem of § 8 that the sheet of solutions is single-valued and that the equations (53) define a one-to-one correspondence between b and β similar to that for a and α.

A similar argument can be made for each of the regions shown in the figure and its corresponding image in the xy-plane.

CHAPTER III

EXISTENCE THEOREMS FOR DIFFERENTIAL EQUATIONS

It is not within the limited scope of these lectures to give a complete account of the various methods for proving the existence of a system of solutions of a set of ordinary differential equations, nor would it be advisable, in view of the many able presentations of these fundamental theorems already well known in mathematical literature. It is rather the intention of the writer to insist on conclusions which can be derived from known methods with regard to the behavior of solutions in any region of size and shape compatible with the continuity properties of the functions by means of which the equations are defined, as over against the usual restriction of the problem to a rectangular or circular neighborhood of a particular point. It has been remarked by Picard* and Painlevé† that if a continuous solution of the differential equation

$$(1) \qquad \frac{dy}{dx} = f(x, y)$$

exists over an interval $\alpha \leqq x \leqq \beta$, then the Cauchy polygons of approximation are defined and converge uniformly to the solution for all values of x in the interval. In § 17 below it is shown that in a region R in which the function f is continuous and satisfies the so-called Lipschitz condition, the polygons of Cauchy passing through a given initial point (ξ, η) interior to R define a priori a continuous solution of the differential equation extending to infinity or else to the boundary of the region. It follows then that there is a function

$$(2) \qquad y = \varphi(x, \xi, \eta)$$

* *Comptes Rendus*, vol. 128 (1899), page 1363.

† *Bulletin de la Société Mathématique de France*, vol. 27 (1899), p. 151.

satisfying the differential equation (1) and defined over a region of points (x, ξ, η) of the form

$$(\xi, \eta) \text{ interior to } R, \quad \alpha(\xi, \eta) < x < \beta(\xi, \eta),$$

and as x approaches α or β the only limiting points which the points (x, y) defined by the function (2) can have are at infinity or else on the boundary of the region R.

In § 18 attention is called to the theorems of Bendixon by means of which it can be shown that the function φ is continuous, and in certain circumstances differentiable with respect to the arguments ξ, η as well as with respect to x. The "imbedding theorem" of Bolza* which asserts that any given solution, near which the function f has suitable continuity properties, can be imbedded in a one-parameter family of neighboring solutions of the differential equation, is an immediate consequence of these results, an analogue for differential equations of the fundamental theorem for implicit functions proved in § 1.

The methods mentioned above are applicable almost without change of wording to a system of equations

$$\frac{dy_\beta}{dx} = f_\beta\,(x, y_1, y_2, \cdots, y_n) \qquad (\beta = 1, 2, \cdots, n)$$

when the symbols y and f in equations (1) are interpreted as row letters in the way apparently first introduced for differential equations by Peano.†

An interesting deduction from the theorems for a system of equations is the proof of the existence of a solution of a partial differential equation

$$F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) = 0$$

which is not necessarily analytic in its five arguments, by means of the well-known theory of characteristic curves, as described in § 19.

* Vorlesungen über Variationsrechnung, page 179.

† "Intégration par séries des équations différentielles linéaires," *Mathematische Annalen*, vol. 32 (1888), p. 450.

§ 15. The Convergence Inequality

There is an inequality which is of frequent service in the existence proof of the following sections and which can be readily deduced from a simple preliminary theorem.

If u is a single-valued function of t with a well-defined forward derivative u' at each point of the interval $0 \leqq t \leqq t_1$, and if

$$|u'| < k|u| + l,$$

k and l being two positive constants, then u also satisfies the inequality

$$|u| \leqq |u_0|e^{kt} + \frac{l}{k}(e^{kt} - 1),$$

where u_0 is the initial value of u at $t = 0$.

Consider the function

$$v = |u_0|e^{kt} + \frac{l}{k}(e^{kt} - 1)$$

satisfying the differential equation

$$v' = kv + l$$

and having $|u_0|$ as its initial value. The value of u is never greater than that of v, since otherwise the difference $u - v$ would vanish and have a positive or vanishing forward derivative at some point. At a point where u and v are equal, however,

$$|u'| < k|u| + l = kv + l = v',$$

which is a contradiction. A similar argument shows that $-u$ is always less than v.

If u is a single-valued function of x with well-defined forward and backward derivatives at each point of an interval $x_0 \leqq x \leqq x_1$, and such that

$$|u'| < k|u| + l,$$

then, for any ξ and x in the interval, u also satisfies the inequality

$$(3) \qquad |u| \leqq |u(\xi)|e^{k|x-\xi|} + \frac{l}{k}(e^{k|x-\xi|} - 1).$$

This may be proved from the preceding paragraphs by putting $t = x - \xi$ for values of x greater than ξ, and $t = -x + \xi$ for values less than ξ.

§ 16. The Cauchy Polygons and their Convergence over a Limited Interval

It is proposed to consider a differential equation (1) for which the function $f(x, y)$ is continuous in the interior of a certain region R of the xy-plane, and such that the quotient

$$\frac{f(x, y') - f(x, y)}{y' - y} \tag{4}$$

is finite when (x, y) and (x, y') lie in any closed region whose points are all interior to R.

A so-called Cauchy polygon for the equation (1) through a point (ξ, η) interior to R is defined by means of equations of the form

$$\begin{aligned} y_1 &= \eta + f(\xi, \eta)(x_1 - \xi), \\ y_2 &= y_1 + f(x_1, y_1)(x_2 - x_1), \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ y &= y_{n-1} + f(x_{n-1}, y_{n-1})(x - x_{n-1}). \end{aligned}$$

The division points

$$\xi < x_1 < x_2 < \cdots$$

may be taken for convenience at equal distances δ from each other. Any value $x > \xi$ will lie on one of the intervals $x_{n-1}x_n$, and the polygon will either be well-defined for all such values, or else there will be a constant β such that for every x in the interval $\xi \leqq x < \beta$ the points of the polygon are interior to R, while for $x = \beta$ the corresponding point (x, y) will be a point of the boundary of R. The polygon defined by the equations above may be denoted by $P_1(x)$, and the analogous one when the division points are distant $\delta/2^{n-1}$ from each other by $P_n(x)$.

A common interval $\xi \leqq x \leqq a$ for two functions $P(x)$, $Q(x)$ with respect to any region R may be defined as one over which

both are interior to R, and one such that on any ordinate of the interval all the points between $(x, P(x))$ and $(x, Q(x))$ are also interior points of R.

Consider now a closed region R_1 interior to R and containing the point (ξ, η), and let m and k be two constants greater respectively than the absolute values of $f(x, y)$ and the quotient (4) in the region R_1. *If $l > 0$ is given in advance, the partitions for any two polygons $P(x)$, $Q(x)$ through (ξ, η) can be taken so small that*

$$(5) \qquad |P(x) - Q(x)| \leq \frac{l}{k}\left(e^{k|x-\xi|} - 1\right)$$

for all values of x in any common interval of $P(x)$ and $Q(x)$ with respect to R_1. For at the point (x, y), where $y = P(x)$, the equation

$$P' = f(x, P) + \{f(x_{n-1}, y_{n-1}) - f(x, P)\} = f(x, P) + \rho$$

is satisfied by the forward and backward derivatives of the polygon P. On account of the continuity of $f(x, y)$ there exists for any l a constant μ such that

$$|x - x'| < \mu, \qquad |y - y'| < \mu$$

imply

$$|f(x, y) - f(x', y')| < l/2$$

whenever the points (x, y) and (x, y') are in R_1. If the subdivisions for $P(x)$ are taken less than μ and μ/m in length, it follows that on the polygon $P(x)$

$$|x - x_{n-1}| < \mu, \qquad |P(x) - y_{n-1}| < m|x - x_{n-1}| < \mu,$$

and hence the absolute value of ρ is less than $l/2$. Similarly $Q(x)$ satisfies an equation

$$Q' = f(x, Q) + \sigma,$$

where $|\sigma| < l/2$, provided that its intervals are less in length than μ and μ/m. The difference $P - Q$ has forward and back-

ward derivatives which satisfy the relations

$$|P' - Q'| \leqq |f(x, P) - f(x, Q)| + |\rho| + |\sigma|$$
$$< k|P - Q| + l,$$

and with the help of the lemma of § 15 the desired inequality follows at once, since P and Q have the same initial value η at $x = \xi$.

If $P(x)$ is a polygon and $Q(x)$ a solution of the differential equation, or if both are solutions, the same theorem evidently holds true, because then the function σ is identically zero, or else both ρ and σ vanish.

The polygons $P_n(x)$ all have a common interval. For take positive constants a and b such that the rectangle

$$(6) \qquad 0 \leqq x - \xi \leqq a, \qquad |y - \eta| \leqq b$$

is entirely within R, and consequently has two constants m and k analogous to those above for R_1. The portions of the polygons in the rectangle (6) all lie between the straight lines

$$y - \eta = \pm m(x - \xi),$$

since the slope of any side of any one of them is numerically less than m. It follows that each is certainly well defined and within the rectangle over an interval $\xi \leqq x \leqq a_1$, where a_1 is the smaller of a and b/m.

The sequence of polynomials $P_n(x)$ converges uniformly, on the interval $\xi \leqq x \leqq \xi + a_1$, to a function $y(x)$ which has a continuous derivative and satisfies the differential equation (1). The curve $y = y(x)$ so defined is entirely within the region R.

For take $\epsilon > 0$ arbitrarily, and l so small that

$$\frac{l}{k}\{e^{ka_1} - 1\} < \epsilon.$$

Then

$$|P_{n'}(x) - P_n(x)| < \frac{l}{k}\{e^{ka_1} - 1\} < \epsilon,$$

provided that the intervals $\delta/2^{n'-1}$ and $\delta/2^{n-1}$ are each less than the constant μ corresponding to l. Hence the sequence $P_n(x)$ converges uniformly to a continuous function $y(x)$ on the interval $\xi \leqq x \leqq \xi + a_1$.

The equations

$$P_n(x) = \eta + \int_\xi^x P_n'(x)dx = \eta + \int_\xi^x \{f(x, P_n) + \rho_n\}dx$$

hold for every n, and the sequences $\{f(x, P_n)\}$ and $\{\rho_n\}$ approach uniformly the limits $f(x, y(x))$ and zero, respectively. Hence

$$y(x) = \eta + \int_\xi^x f(x, y(x))dx;$$

from which it follows by differentiation that $y(x)$ is a solution of the differential equation.

It is easy to show by means of the convergence inequality that there is only one continuous solution $y = y(x)$ of the differential equation (1) in the region R and passing through (ξ, η). For suppose there were another, $Y(x)$, distinct from $y(x)$ at a value $x' > \xi$. There would then be a value $\xi_1 < x'$ at which $y(\xi_1) = Y(\xi_1)$, and such that the two solutions would be distinct throughout the interval $\xi_1 < x \leqq x'$. In a neighborhood of the point of intersection (ξ_1, η_1) interior to R a relation

$$\left|\frac{d(Y - y)}{dx}\right| = |f(x, Y) - f(x, y)| < k|Y - y|$$

would be satisfied, and hence, from the convergence inequality (3),

$$|Y - y| \leqq 0.$$

This contradicts the hypothesis that $y(x)$ and $Y(x)$ are distinct throughout the interval $\xi_1 < x \leqq x'$.

§ 17. The Existence of a Solution Extending to the Boundary of the Region R

It has been proved in the preceding section that, on a certain interval $\xi \leqq x \leqq \xi + a_1$, the polygonal curves $y = P_n(x)$ converge uniformly to a continuous solution $y = y(x)$ of the differential equation (1) lying entirely within the region R. The interval for which the proof has been given may not be the largest one on which the sequence of polygons has this property. There will, however, be a number $\beta \geqq \xi + a_1$, possibly infinity, with the property that on any interval $\xi \leqq x \leqq \beta_1$, where $\beta_1 < \beta$, the sequence of polygons converges uniformly to a continuous solution interior to R. A continuous curve $y = y(x)$ is thus defined which has a derivative and satisfies the differential equation for all values of x in the interval $\xi \leqq x < \beta$.

As x approaches β the points $(x, y(x))$ of the solution can have no limit point (β, γ) interior to the region R.

If they did, there would be for any given ϵ a value $x' < \beta$ such that

$$|x' - \beta| < \epsilon, \qquad |y(x') - \gamma| < \frac{\epsilon}{2},$$

and an integer N such that, whenever $n \geqq N$, the inequality

$$|P_n(x) - y(x)| < \frac{\epsilon}{2}$$

would hold for all values of x in the interval $\xi \leqq x \leqq x'$. At the value x' in particular

$$|P_n(x') - \gamma| \leqq |P_n(x') - y(x')| + |y(x') - \gamma| < \epsilon;$$

so that for $n \geqq N$ the points $(x', P_n(x'))$ would all lie in the ϵ-neighborhood of the point (β, γ). About the point (β, γ) as center a rectangle

$$|x - \beta| \leqq A, \qquad |y - \gamma| \leqq B$$

could be described entirely within the region R, and in the portion

R_1 of R which lay within the rectangle or within the region

$$\xi \leqq x \leqq x', \qquad y(x) - \epsilon < y < y(x) + \epsilon$$

the absolute values of $f(x, y)$ and the quotient (4) would be less than two constants m and k, respectively. It can be shown without great difficulty that every polygon $P_n(x)$ for $n > N$ would be defined and lie within the region R for an interval extending beyond β at least a distance A_1, where A_1 is the smaller of the numbers A and $(B - \epsilon - m\epsilon)/m$. A proof similar to that of § 16 would then show that the polygons $P_n(x)$ converge uniformly to a continuous solution of equation (1) interior to R_1 over an interval $\xi \leqq x \leqq \beta + A_1$; and consequently β could not be the upper bound described above.

As x approaches β, therefore, the only limiting points of the solution $y = y(x)$ are at infinity or else are boundary points of the region R. If R is further a closed region, that is, one containing all of its limit points, then there is but one limit point for the curve $y = y(x)$ as x approaches β. For suppose (β, γ) to be a finite point in any neighborhood of which there are points on the curve. About (β, γ) a rectangle

$$|x - \beta| \leqq A, \qquad |y - \gamma| \leqq B \tag{7}$$

can be chosen arbitrarily, and the points of R lying in it form a finite closed set in which $|f(x, y)|$ remains always less than a constant M. On the interval $\beta - A_1 < x < \beta$, where A_1 is the smaller of the numbers A and B/M, all the points of the curve $y = y(x)$ satisfy the inequality

$$|y - \gamma| \leqq M(\beta - x). \tag{8}$$

For if (x', y') is any point of the curve in the rectangle (7) and also in an ϵ-neighborhood of the point (β, γ), then the inequality

$$\begin{aligned} |y - \gamma| &\leqq |y' - y| + |y' - \gamma| \\ &< M(x' - x) + \epsilon \\ &< M(\beta - x) + \epsilon \end{aligned}$$

must be satisfied by any preceding point $P(x, y)$ of the curve $y = y(x)$ for which the arc PP' is interior to the rectangle. It follows that the solution must lie interior to the rectangle and satisfy the last inequality, at least on an interval $x' - A_\epsilon < x < x'$, where A_ϵ is the smaller of $A - \epsilon$ and $(B - \epsilon)/M$. Hence the inequality (8) is also true on a properly chosen interval preceding $x = \beta$. It follows that as x approaches β there can be but one limit point for the curve $y = y(x)$, and this limit point is either at infinity or else is a boundary point of the region R

When the function $f(x, y)$ in the differential equation

$$\frac{dy}{dx} = f(x, y)$$

satisfies in a region R the conditions stated at the beginning of § 16, *there exists through any interior point (ξ, η) of the region R one and but one continuous solution*

$$y = \varphi(x, \xi, \eta) \tag{9}$$

of the differential equation. This solution is defined and interior to R for all values of x interior to an interval

$$\alpha(\xi, \eta) < x < \beta(\xi, \eta), \tag{10}$$

while as x approaches one of the end values α or β, the only limiting points of the solution are either at infinity or else on the boundary of R. If the region R is closed, then the solution has a unique finite or infinite limit point as x approaches α or β.

§ 18. The Continuity and Differentiability of the Solutions

It can be shown by methods due to Bendixon* that the function $\varphi(x, \xi, \eta)$ and its derivative $\varphi_x(x, \xi, \eta)$, whose existence has been proved in the preceding sections, are continuous in all three of their arguments, and if the function $f(x, y)$ has continuous first derivatives with respect to x and y in the interior of the region R,

* *Bulletin de la Société Mathématique de France*, vol. 24 (1896), p. 220.

then φ and φ_x have also continuous first derivatives with respect to all of their arguments.

The continuity at any set of values (x, ξ, η) for which (ξ, η) is in R and x satisfies the inequality (10) is provable with the help of the convergence inequality of § 15. For there will always be a region R_δ about the arc S of the solution (9) over the

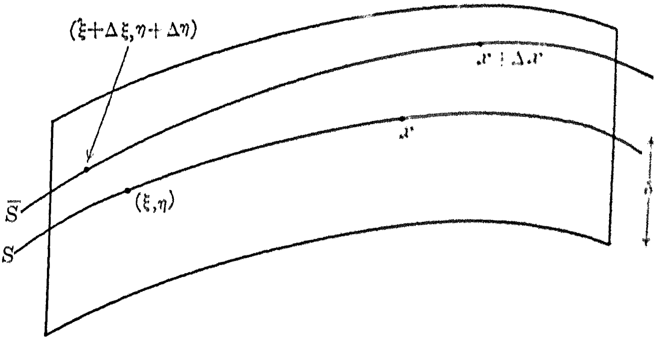

FIG. 11.

interval from ξ to x, of the kind symbolized in the figure, and so small that it lies entirely within the region R. If $(\xi + \Delta\xi, \eta + \Delta\eta)$ is any point in R, then the solution

$$(\bar{S}) \qquad y = \varphi(x, \xi + \Delta\xi, \eta + \Delta\eta)$$

satisfies the inequality

$$(11) \quad \begin{aligned} |\varphi(\xi + \Delta\xi, \xi + \Delta\xi, \eta + \Delta\eta) - \varphi(\xi + \Delta\xi, \xi, \eta)| \\ = |\Delta\eta + \eta - \varphi(\xi + \Delta\xi, \xi, \eta)| \\ \leqq |\Delta\eta| + m|\Delta\xi|, \end{aligned}$$

where m is the maximum of the absolute value of $f(x, y)$ in R_δ, on account of the relation

$$(12) \quad |\eta - \varphi(\xi + \Delta\xi, \xi, \eta)| = \left| \int_{\xi + \Delta\xi}^{\xi} \varphi(x, \xi, \eta)dx \right| \leqq m|\Delta\xi|.$$

Hence as long as $\bar{S}$ remains within the region R_δ, it satisfies the

convergence inequality

$$|\varphi(x,\ \xi + \Delta\xi,\ \eta + \Delta\eta) - \varphi(x,\ \xi,\ \eta)| \leqq \{|\Delta\eta| + m|\Delta\xi|\}e^{|x-\xi-\Delta\xi|},$$

the initial values of the two solutions being taken at $x = \xi + \Delta\xi$. If $\Delta\xi$ and $\Delta\eta$ are sufficiently small the expression on the right is less than δ for all values of x belonging to the region R_δ, and hence $\bar{S}$ must be defined and interior to R_δ for all such values. Otherwise, for some interior value of x, it would attain one of the values $\varphi(x,\ \xi,\ \eta) \pm \delta$, which is seen to be impossible on account of the choice just made of $\Delta\xi$ and $\Delta\eta$.

Consider now the difference

$$\begin{aligned}|\varphi(x+\Delta x,\ \xi+\Delta\xi,\ \eta+\Delta\eta)-\varphi(x,\ \xi,\ \eta)|&\\ \leqq|\varphi(x+\Delta x,\ \xi+\Delta\xi,\ \eta+\Delta\eta)&-\varphi(x,\ \xi+\Delta\xi,\ \eta+\Delta\eta)|\\ &+|\varphi(x,\ \xi+\Delta\xi,\ \eta+\Delta\eta)-\varphi(x,\ \xi,\ \eta)|.\end{aligned}$$

By a step similar to (12), and the inequality (11), it is seen to be less than

$$m|\Delta x| + \{|\Delta\eta| + m|\Delta\xi|\}e^{k|x-\xi-\Delta\xi|}$$

whenever $\Delta\xi$ and $\Delta\eta$ have been so chosen that $\bar{S}$ lies entirely in the region R_δ. Hence the continuity of $\varphi(x,\ \xi,\ \eta)$ is proved.

To prove the differentiability of φ with respect to ξ and η, assume that $f(x,\ y)$ has a continuous derivative f_y in the region R, and consider the same solutions S and $\bar{S}$ in the region R_δ. The difference of their ordinates satisfies the equation

$$\frac{d\Delta\varphi}{dx} = f(x,\ \varphi + \Delta\varphi) - f(x,\ \varphi) = A\Delta\varphi,$$

where, by Taylor's formula with the integral form of remainder,

$$A = \int_0^1 f_y(x,\ \varphi + u\Delta\varphi)du$$

is a continuous function of x, $\Delta\xi$, $\Delta\eta$, the values ξ, η being con-

sidered as constant for the moment. Hence

$$\Delta\varphi = ce^{\int_\xi^x A\,dx}.$$

When $\Delta\xi = 0$ or $\Delta\eta = 0$, the constant c has respectively the values

$$c = \Delta\varphi|_{x=\xi} = \varphi(\xi,\ \xi,\ \eta + \Delta\eta) - \varphi(\xi,\ \xi,\ \eta) = \Delta\eta,$$

$$\begin{aligned} c &= \varphi(\xi,\ \xi + \Delta\xi,\ \eta) - \varphi(\xi,\ \xi,\ \eta) = \int_{\xi+\Delta\xi}^{\xi} f(x,\ \varphi + \Delta\varphi)dx \\ &= -\Delta\xi f(\xi + \theta\Delta\xi,\ \varphi(\xi + \theta\Delta\xi,\ \xi + \Delta\xi,\ \eta)), \end{aligned}$$

where $0 < \theta < 1$. Hence the quotients $\Delta\varphi/\Delta\xi$, $\Delta\varphi/\Delta\eta$ have well-defined limiting values

$$\frac{\partial\varphi}{\partial\xi} = -f(\xi,\ \eta)e^{\int_\xi^x f_\varphi(x,\ \phi)dx}, \qquad \frac{\partial\varphi}{\partial\eta} = e^{\int_\xi^x f_\varphi(x,\ \phi)dx}.$$

It may be remarked in conclusion that the theorems which have been proved in §§ 16–18 are true for systems of equations as well as for a single one.

§ 19. An Existence Theorem for a Partial Differential Equation of the First Order which is not Necessarily Analytic

Proofs have been given by Cauchy, Kowalewski, Darboux, and others for the theorem that in general there exists one and but one analytic surface

$$z = z(x,\ y)$$

which passes through an arbitrarily selected analytic curve C in the xy-space and, with the derivatives

$$p = \frac{\partial z}{\partial x}, \qquad q = \frac{\partial z}{\partial y},$$

satisfies a differential equation of the form

$$F(x,\ y,\ z,\ p,\ q) = 0,$$

where F is an analytic function of its five arguments. These proofs, however, say nothing about the solutions which may exist through a curve C whose defining functions are not expressible by means of power series; and they are not applicable when F itself has not this property. An existence proof is to be given below which is based upon much less restrictive assumptions on the functions F and the curve C. It involves the well-known theory of characteristic strips, which are solutions of a set of ordinary differential equations. If a one-parameter family of characteristic strips intersecting a given curve C is properly selected, it will generate a surface S which is a solution of the differential equation. The existence of the family and the differentiability of the surface depend, however, upon the existence and differentiability of the equations of the characteristic strips with respect to the initial values of the variables which they involve, that is, upon theorems similar to those which have been developed in the preceding sections.

Suppose that the function F is continuous and has continuous first and second derivatives in a certain region R of points (x, y, z, p, q). The differential equations satisfied by the characteristic strips have the form

$$(13)\qquad \begin{aligned} &\frac{dx}{du} = F_p, \quad \frac{dy}{du} = F_q, \quad \frac{dz}{du} = pF_p + qF_q, \\ &\frac{dp}{du} = -F_x - pF_z, \quad \frac{dq}{dx} = -F_y - qF_z. \end{aligned}$$

Through any initial values $(\xi, \eta, \zeta, \pi, \kappa)$ interior to R these equations have a solution with equations and initial conditions of the form

$$(14)\qquad \begin{aligned} x &= x(u, \xi, \eta, \zeta, \pi, \kappa), & \xi &= x(0, \xi, \eta, \zeta, \pi, \kappa), \\ y &= y(u, \xi, \eta, \zeta, \pi, \kappa), & \eta &= y(0, \xi, \eta, \zeta, \pi, \kappa), \\ z &= z(u, \xi, \eta, \zeta, \pi, \kappa), & \zeta &= z(0, \xi, \eta, \zeta, \pi, \kappa), \\ p &= p(u, \xi, \eta, \zeta, \pi, \kappa), & \pi &= p(0, \xi, \eta, \zeta, \pi, \kappa), \\ q &= q(u, \xi, \eta, \zeta, \pi, \kappa), & \kappa &= q(0, \xi, \eta, \zeta, \pi, \kappa), \end{aligned}$$

and such that each of the functions on the left and its derivative for u are continuous and have continuous first derivatives in a region of values $(u, \xi, \eta, \zeta, \pi, \kappa)$ for which $(\xi, \eta, \zeta, \pi, \kappa)$ is a point interior to R and u lies in an interval, containing the value $u = 0$, of the form,

$$\alpha(\xi, \eta, \zeta, \pi, \kappa) < u < \beta(\xi, \eta, \zeta, \pi, \kappa).$$

The points (x, y, z, p, q) so defined are all interior to the region R.

Along the solution (14) the equations

$$(15) \qquad px_u + qy_u - z_u = 0$$

$$(16) \qquad \frac{dF}{du} = F_x x_u + F_y y_u + F_z z_u + F_p p_u + F_q q_u = 0$$

are satisfied identically, so that the direction $p : q : - 1$ is always normal to the curve defined by the first three equations. Evidently if F vanishes at a single point of the strip, it will also vanish at every other point. The solutions (14) along which F vanishes are called characteristic strips, and any one of the strips (14) will surely be of this type if the initial condition

$$F(\xi, \eta, \zeta, \pi, \kappa) = 0$$

is satisfied.

Consider now a continuous and differentiable strip of elements

$$(17) \qquad x = \xi(v), \quad y = \eta(v), \quad z = \zeta(v), \quad p = \pi(v), \quad q = \kappa(v)$$
$$(v_1 \leqq v \leqq v_2)$$

which lies in the interior of the region R and satisfies the conditions

$$(18) \qquad \begin{matrix} \pi\xi_v + \kappa\eta_v - \zeta_v = 0, \\ F(\xi, \eta, \zeta, \pi, \kappa) = 0, \end{matrix} \qquad \begin{vmatrix} F_\pi & \xi_v \\ F_\kappa & \eta_v \end{vmatrix} \neq 0,$$

where the arguments in the derivatives of F are the same as those in the second equation. The first two of these conditions imply that the direction $\pi : \kappa : - 1$ is normal to the curve

$$(19) \qquad x = \xi(v), \quad y = \eta(v), \quad z = \zeta(v),$$

and that the curve and its strip of normals satisfy the differential equation. The third prevents the strip from being a so-called integral strip of the differential equation, through which there does not in general pass a unique integral surface without singularities. To make the situation simpler it will be supposed that the projection of the strip (17) in the xy-plane does not intersect itself.

When the functions (17) are substituted in the equations (14), a new system

$$(20)\qquad \begin{gathered} x = X(u, v), \quad y = Y(u, v), \quad z = Z(u, v), \\ p = P(u, v), \quad q = Q(u, v) \end{gathered}$$

with the initial conditions

$$(21)\qquad \begin{gathered} \xi(v) = X(0, v), \quad \eta(v) = Y(0, v), \quad \zeta(v) = Z(0, v), \\ \pi(v) = P(0, v), \quad \kappa(v) = Q(0, v) \end{gathered}$$

is determined. There is a region

$$(R_{uv})\qquad A \leqq u \leqq B, \quad v_1 \leqq v \leqq v_2,$$

where A is a negative and B a positive constant, in which the functions (20) are continuous, have continuous first derivatives, and satisfy the relation

$$(22)\qquad \begin{vmatrix} X_u & X_v \\ Y_u & Y_v \end{vmatrix} \neq 0.$$

For if M is the maximum of the absolute values of the functions on the right in the equations (13), for a closed ϵ-neighborhood of the points of the strip (17) in the interior of R, then the solutions (14) are defined at least over an interval $|u| \leqq \epsilon/M$, and the absolute values of A and B can be taken at least as great as this constant without disturbing the continuity properties desired for the functions (20) in the region R_{uv}. The condition (22) is satisfied for the values $u = 0$, $v_1 \leqq v \leqq v_2$ because of the first two of equations (13) and the third of the relations (18); and the

region R_{uv} can therefore be chosen so that the determinant is different from zero everywhere in it.

By an argument similar to that used in proving the theorem of § 4 it can be shown that A and B can be restricted still further, if necessary, so that no two distinct points (u', v'), (u'', v'') in the region R_{uv} define the same point (x, y) by means of equations (20). The boundary of the region R_{uv} is transformed then by the first two of equations (20) into a simply closed regular curve in the xy-plane which bounds a portion R_{xy} of the xy-plane. The equations establish furthermore a one-to-one correspondence between the points of R_{uv} and those of R_{xy}, and the functions

$$u = u(x, y), \quad v = v(x, y) \tag{23}$$

so defined are continuous and have continuous first derivatives in R_{xy}. The others of the equations (20) define then three functions

$$z = z(x, y), \quad p = p(x, y), \quad q = q(x, y) \tag{24}$$

which are also continuous and have continuous first derivatives in R_{xy}, and which with the values (23) for u and v satisfy the equations (20) identically in x, y.

The functions (20) satisfy the relations

$$\begin{aligned} PX_u + QY_u - Z_u &= 0, \\ PX_v + QY_v - Z_v &= 0, \\ F(X, Y, Z, P, Q) &= 0, \end{aligned} \tag{25}$$

identically in u, v. The first and third of these follow at once from the equations (15), (16), the second of the equations (18), and (21). The expression

$$\Omega(u, v) = PX_v + QY_v - Z_v$$

has the initial values

$$\Omega(0, v) = \pi\xi_v + \kappa\eta_v - \zeta_v = 0, \tag{26}$$

which vanish on account of the first of equations (18). Furthermore

$$\Omega_u = P_u X_v + Q_u Y_v + P X_{uv} + Q Y_{uv},$$

and from the first of equations (25),

$$0 = P_v X_u + Q_v Y_u + P X_{uv} + Q Y_{uv}.$$

By subtracting the last expression from that for Ω_u and using the equations (13) which the functions (20) satisfy, it follows that

$$\Omega_u = P_u X_v + Q_u Y_v - P_v X_u - Q_v Y_u = -\Omega F_z - \frac{\partial F}{\partial v},$$

in which the arguments of the derivatives of F are the functions (20). Hence with the help of the third of equations (25) and the initial values (26),

$$\Omega_u = -\Omega F_z, \quad \Omega = \Omega(0, v) e^{\int_0^u F_z du'} = 0.$$

The single-valued function $z(x, y)$ defined above over the region R_{xy} has the derivatives

$$Z_x = \frac{\begin{vmatrix} Z_u & Y_u \\ Z_v & Y_v \end{vmatrix}}{\begin{vmatrix} X_u & Y_u \\ X_v & Y_v \end{vmatrix}} = p(x, y), \quad Z_y = \frac{\begin{vmatrix} X_u & Z_u \\ X_v & Z_v \end{vmatrix}}{\begin{vmatrix} X_u & Y_u \\ X_v & Y_v \end{vmatrix}} = q(x, y),$$

found by substituting the functions (23), (24) in the equations (20), differentiating the resulting identities, and applying the first two of the relations (25). It satisfies the differential equation $F = 0$ on account of the third of the equations (25). Furthermore

$$x, \quad y, \quad z(x, y), \quad p(x, y), \quad q(x, y)$$

reduce to $\xi, \eta, \zeta, \pi, \kappa$ at any point of the strip (17), since at such a point $u(\xi, \eta) = 0$ and the relations (21) are satisfied.

It has been proved therefore that there is a single-valued

function

$$z = z(x, y), \tag{27}$$

defined over a region R_{xy} of the xy-plane, which is continuous and has continuous first and second derivatives, contains the initial strip (17), and satisfies the differential equation $F = 0$.

There is no other surface

$$z = z_1(x, y) \tag{28}$$

defined over the region R_{xy} and having these properties. If there were such a one, it would have to contain all of the points of the strips defined by equations (20). To prove this, suppose that (x', y', z', p', q') is an element belonging to one of the strips (20) for values (u', v'), and also to the surface (28). The equations

$$\frac{dx}{du} = F_p(x, y, z_1, p_1, q_1), \quad \frac{dy}{du} = F_q(x, y, z_1, p_1, q_1), \tag{29}$$

where p_1 and q_1 are the derivatives of z_1, have a unique solution

$$x = x_1(u), \quad y = y_1(u) \tag{30}$$

reducing to x', y' for the initial value $u = u'$ and defined over an interval $u' - \epsilon \leq u \leq u' + \epsilon$. The corresponding equations

$$\begin{gathered} x = x_1(u), \quad y = y_1(u), \quad z = z_1(u), \\ p = p_1(u), \quad q = q_1(u), \end{gathered} \tag{31}$$

found by substituting the functions (30) in z_1, p_1, q_1, define a characteristic strip. For on the surface (28) the equations

$$\begin{gathered} F_x + F_z p_1 + F_p r_1 + F_q s_1 = 0, \\ F_y + F_z q_1 + F_p s_1 + F_q t_1 = 0 \end{gathered}$$

are identities in x, y, where r_1, s_1, t_1 are the three second derivatives of $z_1(x, y)$. As a result of these identities and the equations (29),

$$
\begin{aligned}
\frac{dz_1}{du} &= p_1\frac{dx}{du} + q_1\frac{dy}{du} = p_1F_{p_1} + q_1F_{q_1}, \\
\frac{dp_1}{du} &= r_1\frac{dx}{du} + s_1\frac{dy}{du} = - F_x - p_1F_{z_1}, \\
\frac{dq_1}{du} &= s_1\frac{dx}{du} + t_1\frac{dy}{du} = - F_y - q_1F_{z_1},
\end{aligned} \tag{32}
$$

where the arguments of the derivatives of F are the functions (31). The equations (29) and (32) show that the strip (31) is a characteristic strip. Its initial element for $u = u'$ is (x', y', z', p', q'), the same as that for the strip (20) corresponding to $v = v'$. Hence the two must coincide on the interval $u' - \epsilon \leqq u \leqq u' + \epsilon$ on which both are defined.

The initial element (21) of any one of the strips (20) is by hypothesis on the surface (28). According to the last paragraph all of the elements of the strip in an interval $|u| \leqq \epsilon$ must also lie on the surface, and it follows that there can be no upper bound except B for the values of u for which this is true. If $u' < B$ were such a limiting value, the element (x', y', z', p', q') corresponding to u' on the characteristic strip would also belong to the surface, on account of the continuity of $z_1(x, y)$ and its derivatives; and the interval of coincidence would therefore be necessarily longer than $0 \leqq u < u'$.

For any point (x, y) in the region R_{xy} there is but one set of values (u, v) solving the first two of equations (20), and the corresponding value of z from the third equation belongs to both of the surfaces (27) and (28). The two surfaces must therefore coincide throughout.

Suppose now that an initial curve of the form (19) is given instead of the initial strip (17). If to any value v_0 defining a point (ξ_0, η_0, ζ_0) of the curve there corresponds a direction $\pi_0 : \kappa_0 : - 1$ satisfying the relations (18), and such that $(\xi_0, \eta_0, \zeta_0, \pi_0, \kappa_0)$ is interior to R, then there will be a strip of elements of the form (17) along the curve containing these initial values for $v = v_0$. For the first two equations (18) have the solution (v_0, π_0, κ_0)

when their first members are regarded as functions of v, π, κ, and on account of the third relation (18) their functional determinant for π, κ does not vanish at these values. According to the fundamental theorem of § 1 there is therefore a pair of functions $\pi(v)$, $\kappa(v)$ defined over an interval $v_1 < v < v_2$ containing v_0 and satisfying, with $\xi(v)$, $\eta(v)$, $\zeta(v)$, the relations (18).

The results of the preceding paragraphs may be summarized as follows:

Suppose that

$$(C) \qquad x = \xi(v), \quad y = \eta(v), \quad z = \zeta(v)$$

is a continuous and differentiable curve, at some point $(\xi_0, \eta_0, \zeta_0) = (\xi(v_0), \eta(v_0), \zeta(v_0))$ *of which there is a normal* $\pi_0 : \kappa_0 : -1$ *satisfying the equation*

$$F(\xi_0, \eta_0, \zeta_0, \pi_0, \kappa_0) = 0.$$

Suppose furthermore that

$$\begin{vmatrix} \xi_v(v_0) & F_p(\xi_0, \eta_0, \zeta_0, \pi_0, \kappa_0) \\ \eta_v(v_0) & F_q(\xi_0, \eta_0, \zeta_0, \pi_0, \kappa_0) \end{vmatrix} \neq 0,$$

and that the initial element $(\xi_0, \eta_0, \zeta_0, \pi_0, \kappa_0)$ *lies in a region* R *of points* (x, y, z, p, q) *in which* F *is continuous and has continuous first and second derivatives. Then there is a strip of the form*

$$(S) \qquad x = \xi(v), \quad y = \eta(v), \quad z = \zeta(v), \quad p = \pi(v), \quad q = \kappa(v)$$
$$(v_1 \leqq v \leqq v_2)$$

containing $(\xi_0, \eta_0, \zeta_0, \pi_0, \kappa_0)$ *for* $v = v_0$, *and such that all of its elements have the properties ascribed above to this initial one. If the projection* C_{xy} *of* C *in the xy-plane does not intersect itself, the characteristic strips of the differential equation*

$$F(x, y, z, p, q) = 0$$

which pass through the elements of S *simply cover a region* R_{xy} *of*

the xy-plane and envelop a single-valued surface

$$z = z(x, y).$$

This surface is continuous and has continuous first and se derivatives in R_{xy}, contains the strip S, and satisfies the differe equation $F = 0$. There is no other surface over the region R_{xy} t has these properties.

DIFFERENTIAL-GEOMETRIC ASPECTS OF DYNAMICS

BY

EDWARD KASNER

CONTENTS

CHAPTER III

Transformation Theories in Dynamics

CHAPTER IV

Constrained Motions in a Field. Generalization of the Trajectory Problem Including Brachistochrones and Catenaries

CHAPTER V

More Complicated Types of Force

DIFFERENTIAL-GEOMETRIC ASPECTS OF DYNAMICS

BY

EDWARD KASNER

INTRODUCTION

The relations between mathematics and physics have been presented so frequently and so adequately in recent years, that further discussion would seem unnecessary. Mathematics, however, is too often taken to be analysis, and the role of geometry is neglected. Geometry may be viewed either as a branch of pure mathematics, or as the simplest of the physical sciences. For our discussion we choose the latter point of view: geometry is the science of actual physical or intuitive space. All physical phenomena take place in space, and hence necessarily present geometric aspects. We confine our discussion to mechanics, and consider the rôle of geometry in mechanics.

The fundamental concepts of mechanics are: space, time, mass, and force. Certain preliminary theories deal with some instead of all these concepts. Space by itself gives rise to pure geometry with all its subdivisions. According to Sir William Rowan Hamilton, algebra is the science of pure time; in fact time is the simplest one-dimensional manifold suggesting the notion of real number, the continuum, the foundation of analysis. Neither mass by itself, nor force by itself, gives rise to an independent theory, for these notions cannot be considered without considering space also.

Space and time together give rise to kinematics. If we do not consider velocities and accelerations, but only displacements (that is, initial and terminal positions without introducing

continuous motion from one to the other), we obtain Ampère's "geometry of motion," which belongs to pure geometry rather than to kinematics.

Space and mass give rise to a separate discipline which may be called the geometry of masses. This deals with centers of gravity, moments of inertia, and moments of higher type, which have been studied extensively in recent years, especially by the Italian mathematicians.

Space and force are the essential concepts employed in rigid statics. Mass and time are not necessary in this theory, which deals essentially with the equivalence and reduction of systems of vectors. The remaining combinations, mass and time, force and time, mass and force, do not produce separate theories, since they can not be discussed without introducing also the concept of space.

Consider then space, time, and mass. The principal development along this line is Hertz's remarkable "geometry and kinematics of material systems," a theory entirely independent of the concept of force.

The other combinations of three of the four concepts have not produced separate developments.

Finally, we have the theory which involves all four concepts simultaneously, namely, kinetics.

Although the geometric aspects of the preliminary theories are very interesting and important, it is not our intention to review the progress which has been made in this line. We mention only Ball's theory of screws, Study's Geometrie der Dynamen, and the law of duality connecting kinematics and statics—a law which is not dynamical, but purely geometric.

The notion of vector is of course fundamental in many of these theories. We recall the fact that there are three distinct types of vector used in mechanics: the *free* vector, the *sliding* vector, the *bound* vector. These three types differ with respect to the definition of equivalence. In the first theory, two vectors are regarded as equivalent when they have the same length

and direction (including sense). Such free vectors are employed in combining translations, or forces acting at a point. A free vector in space has three coordinates. The sum of any number of vectors is a vector.

In the second theory, dealing with sliding vectors, two vectors are equivalent only when they have the same line as well as the same length and sense. Such vectors are used in the statics of rigid bodies. The sliding vector in space has five coordinates. A system of these vectors can not usually be reduced to a single vector. The most general system depends in fact on six essential parameters: it is a new geometric element which may be represented either as a screw or a dyname.

Finally, in the third type of vector theory, two vectors are not equivalent unless they have the same initial point and same terminal point, that is the vector is completely bound. Such a vector in space depends on six coordinates. The most general system depends on twelve essential parameters. This is the theory required in the developments of *astatics*.

Statics and kinematics have given rise to very extensive geometric developments; but kinetics still is thought of almost exclusively as a matter of differential equations. Lagrange, in the famous preface to his Mécanique Analytique, stated that no diagrams would be found in his work: "Lovers of analysis will thank me for adding a new branch to that science." The special object of these lectures will be to point out some of the geometric aspects of kinetics, especially properties of the trajectories described in arbitrary fields of force. While the investigations connected with statics and kinematics are mainly of algebraic-geometric character, our kinetic discussions relate to infinitesimal properties, tangents, distribution of curvature, osculating conics, and so on: we shall deal chiefly with the *differential geometry of systems of trajectories*. It is essential to observe that the properties considered relate not to the individual curves, but to the infinite systems of curves.

To emphasize this point, consider the motion of a particle

in a plane field of force, the force depending only on the position of the point. For given initial conditions, the particle will move on a definite curve; taking all possible initial conditions, we shall obtain a triply infinite system of curves. A single curve obviously has no peculiarities, for a particle may be made to describe any given curve by selecting a proper force, varying from point to point of that curve. The system of curves, however, will have intrinsic peculiarities, for if a triply infinite system of curves is given at random, it will not usually be possible to find any field of force such that every particle moving in that field will describe one of the given curves; there is, for instance, no field of force which produces as its trajectories all the circles of the plane.

The simplest general property of the system of trajectories is as follows: If a particle is started at a given position in a given direction with all possible initial speeds into a field of force, a single infinity of trajectories will be obtained, one for each value of the speed; construct for each of these curves the parabola having four-point contact (osculating parabola); the foci of these parabolas will always lie on a circle passing through the given initial point. An equivalent statement is that the directrices of these parabolas will always be concurrent. In space we employ osculating spheres and find that the locus of the centers is a straight line.

A completely characteristic set of properties, for both the plane and space, is given in Chapter I. It is thus possible to tell when a given system of curves can serve as a system of dynamical trajectories. A method is obtained for constructing the field from its trajectories. If say a handful of particles is thrown into an unknown field (the force acting at any point depending only on the position of the point) and if a photograph of the totality of paths is taken, then, without any record of velocity or any observation of time, the field can be constructed. In particular it is possible, by simple geometric tests, to distinguish conservative from non-conservative fields.

Chapter II deals with the geometry of conservative forces. Here the energy equation allows us to group the trajectories into "natural families." Such a family is obtained most concretely as the totality of ∞^4 rays or paths of light in any medium where the index of refraction varies continuously from point to point. The geometric characterization is first given by two simple properties relating to circles of curvature; and then by a new converse of the theorem of Thomson and Tait. It is seen, for example, that if a candle is placed in the atmosphere or in any gas of variable density, the ∞^2 rays emitted by it, which may be curves of very complicated shape, will necessarily have these properties: (A) the circles of curvature constructed at the given source all meet at a second point; (B) three of these circles have four-point (instead of merely three-point) contact with their curves, and these three are mutually orthogonal; (C) the ∞^2 rays form a normal congruence, that is, admit ∞^1 orthogonal surfaces. Natural families are characterized either by (A) and (B), or by (A) and (C).

These results are applied to the propagation of waves in any isotropic medium. A second and more complicated converse question suggested by the Thomson-Tait theorem is discussed. Some interesting optical theorems are given a geometric formulation, but the converse problems are left unsettled. The final section deals with the "general problem of dynamics" in the sense of the French writers.

The third chapter deals with transformation theories. It is interesting to notice how the most important groups of geometry, the projective and the conformal, play essential rôles in dynamics, the former in connection with arbitrary fields, the latter in connection with conservative fields and natural families. The infinitesimal contact transformations of mechanics, and a new group of space-time transformations are also discussed.

The chief subject of Chapter IV is a simple problem in constrained motion, which includes, and hence serves to unify, the theories of trajectories, brachistochrones, catenaries, and velocity

curves in an arbitrary field of force. Complete characterizations are given. Curves of constant pressure and tautochrones are treated only briefly.

Chapter V includes brief discussions of more complicated problems in motion, for example, the effect of a resisting medium on the geometric character of the system of trajectories; the motions of any number of interacting particles (the results being of course applicable to the problem of three bodies); finally, forces depending not only on position but also on the time, both trajectories and space-time curves being studied. The latter are constructed, in the sense of Minkowski, in the four-dimensional space (x, y, z, t), but the application made is to ordinary dynamics, not to electrodynamics or relativity theory.

The main results of the first two chapters (in particular the complete characterizations of general systems of trajectories and of natural families) were first given by the writer in a series of four papers published in the *Transactions* of this Society (1906–1910). Some of the other results are given in notes published in the *Bulletin*. The last two chapters, as well as many sections of the other chapters, deal with hitherto unpublished results.

CHAPTER I

TRAJECTORIES IN AN ARBITRARY FIELD OF FORCE

§§ 1–8. Trajectories in the Plane

1. We consider first the motion of a particle in the plane under the action of any positional field of force. The general equations of motion are

$$m\frac{d^2x}{dt^2} = \varphi(x, y), \quad m\frac{d^2y}{dt^2} = \psi(x, y),$$

where m is the mass and φ, ψ are the rectangular components of the force acting at any point x, y. There is no loss of generality in assuming the mass of the particle to be unity, so we write*

$$(1) \qquad \ddot{x} = \varphi(x, y), \qquad \ddot{y} = \psi(x, y).$$

The particle may be started from any position (x_0, y_0) with any velocity $(\dot{x}_0, \dot{y}_0)$. A definite trajectory is then described. Since the same curve may be obtained by starting from any one of its ∞^1 points, the total number of trajectories, for all initial conditions, is ∞^3. The differential equation of the third order representing this system of trajectories, found by eliminating the time from (1), is

$$(2) \quad (\psi - y'\varphi)y''' = \{\psi_x + (\psi_y - \varphi_x)y' - \varphi_y y'^2\}y'' - 3\varphi y''^2.$$

This is not an arbitrary differential equation of the third order. Hence the system of trajectories generated by an arbitrary field of force must have peculiar geometric properties, which translate the peculiar analytic form of (2).

* *The following notation is employed throughout these lectures:* Dots indicate total differentiation with respect to the time t; primes indicate total differentiation with respect to x; subscripts x and y indicate partial differentiation; finally, the subscript s indicates total differentiation with respect to the arc length s.

2. Before stating these, we remark that a more intrinsic basis for the discussion is obtained by decomposing the acting force into components normal and tangential to the path, instead of parallel to x and y axis as in (1). Denoting these components by N and T respectively, the equations of motion are

$$v^2/r = N, \quad vv_s = T, \tag{3}$$

where v denotes the speed, s the arc length, and r the radius of curvature. By differentiating the first of these equations with respect to s, and comparing with the second equation, we can eliminate v, obtaining

$$(rN)_s = 2T, \tag{4}$$

a relation which defines the trajectories and is equivalent to (2).

To reduce this to a more explicit form, we introduce an auxiliary vector, completely determined by the given field of force, namely the space derivative of the force (considered of course as a vector). The normal and tangential components of the force vector are

$$N = \frac{\psi - y'\varphi}{\sqrt{1 + y'^2}}, \quad T = \frac{\varphi + y'\psi}{\sqrt{1 + y'^2}}; \tag{5}$$

the corresponding components of the new vector are

$$\begin{aligned} \mathfrak{N} &= \frac{\psi_s - y'\varphi_s}{\sqrt{1 + y'^2}} = \frac{\psi_x + (\psi_y - \varphi_x)y' - \varphi_y y'^2}{1 + y'^2}, \\ \mathfrak{T} &= \frac{\varphi_s + y'\psi_s}{\sqrt{1 + y'^2}} = \frac{\varphi_x + (\varphi_y + \psi_x)y' + \psi_y y'^2}{1 + y'^2}. \end{aligned} \tag{6}$$

While the new vector is the s derivative of the force vector, its components are obviously not the same as the s derivatives of the old components: the correct relations are found to be

$$N_s = \mathfrak{N} - \frac{T}{r}, \quad T_s = \mathfrak{T} + \frac{N}{r}. \tag{7}$$

These formulas are sufficient for the discussion of trajectories.* By means of (7) we can reduce (4) to the form

$$(8) \qquad N r_s = -\mathfrak{N} r + 3T.$$

This is the fundamental intrinsic representation of the system of ∞^3 *trajectories connected with a given field of force.*

From it we may derive very simply a number of geometric properties. But in dealing with the converse questions which arise, and in proving the completeness of the set obtained, it is more convenient to use the equivalent cartesian representation, that is, equation (2).†

3. *The Five Characteristic Properties in the Plane.*—The system of trajectories generated by any positional field of force in the plane has the following set of properties, and conversely, any system of ∞^3 curves which has these properties will be a system of dynamical trajectories.

* In some of the later discussions we shall need also the space derivatives of $\mathfrak{N}$ and $\mathfrak{T}$, which may be written in the form

$$\mathfrak{N}_s = \mathfrak{N}_1 + \frac{\mathfrak{N}_2}{r}, \qquad \mathfrak{T}_s = \mathfrak{T}_1 + \frac{\mathfrak{T}_2}{r},$$

where

$$\mathfrak{N}_1 = \frac{\psi_{xx} + (2\psi_{xy} - \phi_{xx})y' + (\psi_{yy} - 2\phi_{xy})y'^2 - \phi_{yy}y'^3}{(1+y'^2)^{3/2}},$$

$$\mathfrak{N}_2 = \frac{\psi_y - \phi_x - 2(\phi_y + \psi_x)y' + (\phi_x - \psi_y)y'^2}{1+y'^2},$$

$$\mathfrak{T}_1 = \frac{\phi_{xx} + (2\phi_{xy} + \psi_{xx})y' + (\phi_{yy} + 2\psi_{xy})y'^2 + \psi_{yy}y'^3}{(1+y'^2)^{3/2}},$$

$$\mathfrak{T}_2 = \frac{\phi_y + \psi_x + 2(\psi_y - \phi_x)y' - (\phi_y + \psi_x)y'^2}{1+y'^2}.$$

The functions ϕ, ψ depend only on the position of the particle; the auxiliary intrinsic functions N, T, $\mathfrak{N}$, $\mathfrak{T}$, $\mathfrak{N}_1$, $\mathfrak{N}_2$, $\mathfrak{T}_1$, $\mathfrak{T}_2$, defined above, depend also upon the direction of motion; finally, N_s, T_s, $\mathfrak{N}_s$, $\mathfrak{T}_s$ depend upon the curvature of the path. Cf. *Bull. Amer. Math. Soc.*, vol. 15 (1909), p. 475.

† Cf. *Trans. Amer. Math. Soc.*, vol. 7 (1906), pp. 401–424. The result contained in property IV of § 3 gives this simple, but apparently overlooked, dynamical theorem: If a particle starts from rest, the initial curvature of the path described is one third of the curvature of the line of force through the initial position.

I. If for each of the ∞^1 trajectories passing through a given point in a given direction we construct the osculating parabola, at the given point, the locus of the foci of these parabolas is a circle passing through that point.

II. The circle that corresponds, according to property I, to a lineal element, is so situated that the element bisects the angle between the tangent to the circle and a certain direction fixed for the given point (the direction of the force acting at the given point).

III. In each direction at a given point there is one trajectory which has four-point contact with its circle of curvature; the locus of the centers of the ∞^1 hyperosculating circles constructed at the given point is a conic passing through that point in the fixed direction described in property II.

IV. With any point O there is associated a certain conic passing through it as described in property III. The normal to the conic at O cuts the conic again at a distance equal to three times the radius of curvature of the line of force passing through O. (The lines of force are defined geometrically by the fact that the tangent at any point has the direction associated with that point in accordance with property II.)

V. When the point O is moved, the associated conic referred to above changes in the following manner. Take any two fixed perpendicular directions for the x direction and the y direction; through O draw lines in these directions meeting the conic again at A and B respectively. Also construct the normal at O meeting the conic again at N. At A draw a line in the y direction meeting this normal in some point A', and at B draw a line in the x direction meeting the normal in some point B'. The variation property referred to takes the form

$$\frac{\partial}{\partial x}\frac{1}{AA'}+\frac{\partial}{\partial y}\frac{1}{BB'}+\frac{\omega\omega_{xy}-\omega_x\omega_y}{3\omega^2}=0,$$

where AA' and BB' denote distances between points, and where ω denotes the slope of the lines of force relative to the chosen

x direction. This is true for any pair of orthogonal directions,

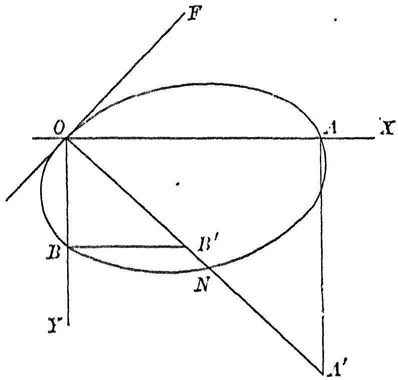

FIG. 1.

and therefore really expresses an intrinsic property of the system of curves.

4. The most general system of ∞^3 curves in the plane is represented by an arbitrary differential equation of the third order

$$(F_0) \qquad y''' = f(x, y, y', y'').$$

It thus involves *one* arbitrary function of *four* arguments.

A system of dynamical trajectories, on the other hand, is represented by an equation of the particular form

$$(F_V) \qquad (\psi - y'\varphi)y''' = \{\psi_x + (\psi_y - \varphi_y)y' - \varphi_y y'^2\}y'' - 3\varphi y''^2,$$

and thus involves *two* arbitrary functions of *two* arguments. These are the only systems having all five properties I–V.

It is interesting to notice just how the successive imposition of the properties gradually narrows down the general form (F_0) to the particular form (F_V).

5. *The most general system having property I is found to be*

$$(F) \qquad y''' = G(x, y, y')y'' + H(x, y, y')y''^2.$$

It thus involves *two* arbitrary functions of *three* arguments. This type of course includes the dynamical type as a very special case. It arises in a number of different geometric and physical investigations. It has therefore its own interest. The characteristic property may be stated in various ways, all of course equivalent to the original form: (I) The osculating parabolas of the trajectories passing through a given point in a given direction have the foci situated on a circle passing through the given point. Five equivalent forms are as follows:

I (2). The directrices of the osculating parabolas form a pencil It follows that there exists a point (the vertex of this pencil) from which all the parabolas subtend an angle of 90°.

I (3). If for each of the trajectories considered, we construct the center of curvature of its evolute, the locus of the centers thus obtained is a parabola passing through O, and having its axis parallel to the given initial direction.

I (4). For each of the trajectories, construct the osculating equiangular spiral. The locus of the centers of the poles of these spirals is a circle passing through O.

I (5). Construct for each of the trajectories the axis of deviation, that is the line bisecting the chords of the curve which are parallel and infinitesimally close to the tangent. The tangent of the angle between the varying axis of deviation and the fixed normal is a linear function of the radius of curvature.

I (6). The derivative of the radius of curvature with respect to the arc length is a linear integral function of the radius of curvature. This is practically a restatement of (5), since for any curve the derivative of the radius of curvature is known to be equal to three times the tangent of the angle of deviation. But in this form it has the advantage of being valid, not only in the plane, but in space of three and in fact any number of dimensions.

If in addition to property I, we impose property II, the function $H(x, y, y')$ is specialized to

$$H = \frac{3}{y' - \omega(x, y)},$$

Thus the most general system with properties I and II is

$$(F_{\mathrm{II}}) \qquad (y' - \omega)y''' = (y' - \omega)Gy'' + 3y''^2$$

where G is any function of x, y, y', and ω is any function of x, y. The type thus involves *one* arbitrary function of *three* arguments and one arbitrary function of two arguments.

6. *Systems with Properties I, II, III.*—Imposing also property III, we find that $G(x, y, y')$ must be of the special form

$$G = \frac{\lambda y'^2 + \mu y' + \nu}{y' - \omega}.$$

Thus the most general system of ∞^3 curves having properties I, II, III is represented by

$$(F_{\mathrm{III}}) \qquad (y' - \omega)y''' = (\lambda y'^2 + \mu y' + \nu)y'' + 3y''^2,$$

involving *four* arbitrary functions ω, λ, μ, ν each of the *two* arguments x, y.

This type may be characterized by the following properties which are then equivalent to I, II, III.

I (2). For a given lineal element, the directrices of the ∞^1 osculating parabolas pass through a common point D.

II (2). When the lineal element turns about the given point O, the point D describes a straight line passing through O.*

III (2). The correspondence between the range of points D and the pencil of elements through O is one-to-two of the special form

$$\frac{3}{2d} = \lambda \sin^2 \theta + \mu \sin \theta \cos \theta + \nu \cos^2 \theta,$$

where d denotes the distance OD, and θ is the angle between the element and fixed direction of OD.

* In the dynamical case this line OD is perpendicular to the force vector acting at O. For certain special fields the point D may remain fixed: this happens only when the components of the force are conjugate harmonic functions, that is when the field is of the type termed "analytic" by Lecornu.

7. If now we add the properties IV and V, the four functions ω, λ, μ, ν appearing in (F_{III}) must obey the relations

$$(F_{\text{IV}}) \qquad \lambda\omega^2 + \mu\omega + \nu + \omega_x + \omega\omega_y = 0,$$
$$(F_{\text{V}}) \qquad (\omega_y + \lambda\omega + \mu)_y - \lambda_x = 0.$$

Thus the general system having properties I–IV involves *three* arbitrary functions of x, y; while that having all five properties involves *two* such functions.

By integrating these relations, we may express the four functions in terms of two arbitrary functions φ, ψ as follows:

$$\omega = \frac{\psi}{\varphi}, \qquad \lambda = \frac{\varphi_y}{\varphi}, \qquad \mu = \frac{\varphi_x - \psi_y}{\varphi}, \qquad \nu = -\frac{\psi_x}{\varphi}.$$

These values, substituted in the type (F_{III}), actually give rise to the type

$$(\psi - y'\varphi)y''' = \{\psi_x + (\psi_y - \varphi_y)y' - \varphi_y y'^2\}y'' - 3\varphi y''^2,$$

and thus the proof is completed that the set of five properties characterizes the dynamical type.

In connection with the statements I (2), II (2), III (2), property IV may be formulated as follows:

IV (2). In the correspondence described in III (2), if the element approaches the direction of the force the corresponding distance OD has for its limiting value 3/2 the radius of curvature of the line of force passing through O. It is to be remembered that the direction of the force, and hence also the lines of force, are defined purely geometrically in terms of the given triply infinite system of curves by the fact that at any point O in the plane the "direction of the force" is perpendicular to the line described as the locus of D in the above equivalent II (2) of property II.

In the same line of ideas it would be possible to find an equivalent for property V (thus completing the characterization), but the result V (2) cannot be put into simple form. The original form V may be criticized as inelegant because in it reference is made to a system of cartesian axes. Of course the result

expresses an intrinsic property since it is true for all systems of axes. It would certainly be desirable to restate the result in intrinsic language. This can be done, for instance, by introducing the distances cut off by the conic described in IV, not only on the normal ON, but also on the two lines inclined at an angle of 45° to that normal. However it does not seem possible to obtain a statement which is both simple and intrinsic in form.

8. Of course many other properties of trajectories may be obtained, either by reasoning synthetically from the five fundamental properties, or by reasoning analytically from the fundamental differential equation. We state only a few samples.

If we shoot particles from a given position in a given direction with variable speed, the center of curvature of the resulting trajectories describes a straight line (the normal) and the focus of the osculating parabola simultaneously describes a circle (by property II), in such a way that the two ranges (one linear, the other circular) are homographically related; furthermore the given point, which is on both ranges, corresponds to itself.

If we shoot from the same position in a direction perpendicular to that previously employed, the new focal circle will be tangent (at the given point) to the former focal circle. Conversely if two focal circles, for the same point, are tangent, the initial directions to which they correspond will be perpendicular to each other.

We shall make use of the following properties which describe the disposition of the ∞^1 focal circles constructed at a given point. The two results which follow are geometrically equivalent, and either may be substituted for property III in the fundamental set.

III (3). If for each of the elements at a given point we construct the corresponding focal circle, the locus of the centers of the ∞^1 circles thus obtained is a conic with one focus at that point.

III (4). The envelope of the ∞^1 focal circles is always a circle.*

* This enveloping circle is in general position: it does not usually have its center at the given point. This simple position arises only when the force is of the Lecornu type.

§ 9. Actual and Virtual Trajectories

9. If we consider the motion of a cannon ball in a given vertical plane under the action of gravity assumed constant, the triply infinite system of trajectories consists of parabolas with vertical axes. We do not, however, obtain all vertical parabolas, represented by the differential equation of the system of trajectories, which is here $y''' = 0$, but only those whose concavity is directed downwards. The other vertical parabolas, with concavity directed upwards, satisfy the same differential equation, and it is therefore convenient to include them in the system studied. We thus have a distinction of *actual* and *virtual* trajectories. The latter are the actual trajectories corresponding to gravity reversed in direction.

In an arbitrary field of force the same distinction arises. The complete system of trajectories is composed of the actual trajectories corresponding to the given force, and the virtual trajectories which are the actual trajectories corresponding to the reversed field. It is obvious that the system of trajectories is not changed if the force acting at every point is multiplied by a constant. If we were considering only actual trajectories, it would be necessary to restrict this constant to positive values, but as we include both actual and virtual, the constant factor may also be negative. (Of course the constant must not be zero, since then the force would vanish and we should obtain the trivial system of straight lines.)

It is easy to show that the virtual trajectories corresponding to the given field may be found by giving the initial speed of the particle a pure imaginary value. The cannon ball could be made to describe a parabola with its concavity directed upwards if only some kind of powder could be invented which would cause its initial speed to be imaginary!

In discussing the general geometric properties of trajectories, we had in mind of course the complete system as defined by the differential equation. Consider for example property I: for any

given lineal element the locus of the foci of the parabolas osculating the corresponding trajectories is a circle through the given point. The question arises, what part of this circle corresponds to the actual trajectories. It is easily found to be the arc of the circle cut off by the initial direction line (the common tangent of the trajectories considered) on that side which is indicated by the force vector. Thus, if we confined our discussion to actual trajectories, the focal locus would be, not a *circle*, but an *arc of a circle*, the arc running from the given point O to a certain terminal point A. If we consider all elements through O the locus of the corresponding terminal point A is found to be a conic passing through O in the direction of the force vector.*

For a given element, the point A, which separates the actual from the virtual, may be defined as the limiting position of the focus of the osculating parabola as the initial speed becomes infinitely large. The osculating parabola in this limiting case becomes a straight line, but the focus has a definite limiting position.

An analogous distinction, into actual and virtual, presents itself also in the theories of brachistochrones, catenaries, and tautochrones. The differential equations of the systems of curves are satisfied by both the actual and virtual curves, and it is the complete systems that we refer to in all our discussions unless the contrary is explicitly mentioned.

§§ 10–15. Trajectories in Space

10. Consider the motion of a particle, which we may take to be of unit mass, in an arbitrary positional field of force. The equations of motion are

$$(1) \qquad \ddot{x} = \varphi(x, y, z), \qquad \ddot{y} = \psi(x, y, z), \qquad \ddot{z} = \chi(x, y, z).$$

The particle may be started from any position, in any direction, with any speed: its motion is then determined by the field of

* This conic is not the same as the conic arising in property III.

force, and it describes a definite trajectory. The totality of trajectories constitutes a definite system of ∞^5 curves. (We exclude the case where the force vanishes at every point, the trajectories then being merely the ∞^4 straight lines.)

What are the properties of such quintuply infinite systems of curves? Obviously an arbitrary system of space curves cannot be obtained as the totality of trajectories connected with any field of force. In fact the most general system of ∞^5 curves (assuming that ∞^1 curves pass through any point of space in any direction) would be represented by a pair of differential equations, one of the third order and one of the second order, of the general form

$$(2)\quad y''' = f(x, y, z, y', z', y''), \qquad z'' = g(x, y, z, y', z', y''),$$

thus involving two arbitrary functions of *six* arguments; while the dynamical type involves merely three arbitrary functions of *three* arguments. The differential equations representing the dynamical type, obtained by eliminating the time from the equations of motion, may be written in the form

$$(3)\quad \begin{aligned} (\psi - y'\varphi)y''' &= \begin{vmatrix} 1 & \varphi_x + y'\varphi_y + z'\varphi_z \\ y' & \psi_x + y'\psi_y + z'\psi_z \end{vmatrix} y'' - 3\varphi y''^2, \\ (\psi - y'\varphi)z'' &= (\chi - z'\varphi)y''. \end{aligned}$$

The question is to express the peculiar form of these equations in simple geometric language.

The interpretation of the second equation is obvious: the osculating plane of the path passes not only through the given initial direction $1 : y' : z'$, but also through the fixed direction $\varphi : \psi : \chi$; that is, the osculating plane always passes through the direction of the force acting at the given point. The other properties are not obvious;* they take into account the form of the differential equation of third order.

* The simplest of these, property II below and certain consequences, were first stated in the author's note published in the *Bull. Amer. Math. Soc.*,

We cannot now, as in the case of the plane discussion, employ osculating parabolas, since our curves are twisted. Three consecutive points of a curve determine an osculating circle. What do four consecutive points determine? No simple type of osculating curve is available, so we shall make use of the osculating sphere. The results are therefore quite different in form from those obtained in the two-dimensional theory.

11. *The Four Properties in Space.*—In order that a system of ∞^5 space curves, of which ∞^1 pass through each point in each direction, shall be identifiable with the system of trajectories generated by a positional field of force, it is necessary and sufficient that it shall have the following four purely geometric properties:

I. The osculating planes of the ∞^3 curves passing through a given point form a pencil; that is, all the planes pass through a fixed direction.

II. The osculating spheres of the ∞^1 curves passing through a given point in a given direction form a pencil; their centers thus lie on a straight line.

III. The straight lines which correspond, in accordance with II, to all the ∞^2 directions at a given point, form a congruence (of order one and of class three) consisting of the secants of a twisted cubic curve; which cubic furthermore passes through the given point in the direction fixed by property I.

IV. The associated plane systems S', determined by the given space system in the manner described below, have the five geometric properties characteristic of a system of plane dynamical trajectories. Consider the given system of ∞^5 space curves in connection with any plane p. Through any point of p there pass ∞^2 curves of the given system which are tangent to the plane. Project the differential elements of the third order belonging to these space curves orthogonally upon p, thus obtaining ∞^2

vol. 12 (1905), pp. 71–74. Somewhat simplified proofs were then given by Cesàro, in a paper published shortly before his death, in the *Memorie di Torino* (1905). The complete set of properties appeared in the *Trans. Amer. Math. Soc.*, vol. 8 (1907), pp. 121–140.

plane differential elements of the third order at the selected point. Applying this process to all points of p, we have a defined set of ∞^4 differential elements of the third order. These elements define a certain differential equation of the third order, and thus determine a system of ∞^3 integral curves. This we term the associated system in the plane p. The space system has the property that every one of these plane systems associated with it is a system of dynamical trajectories, and therefore has the five properties stated in § 3, which we here denote by I_p, V_p in order to avoid confusion with the four spatial properties.

These four properties are ordinally independent: no one can be derived from those which precede it. The question of absolute independence is left open: it is quite probable that IV is sufficiently strong to furnish a complete characterization by itself.

12. The most general system having property I involves one arbitrary function of six arguments besides two functions of three arguments. These systems have the following properties, which are of course consequences of property I.

The ∞^1 curves passing through a given point in a given direction have not only the same osculating plane, but also the same *torsion*.

If the torsion is given the corresponding initial directions form a quadric cone. In particular such a cone defines the directions of those curves, through the given point, which admit hyperosculating planes.

If for each of these curves we construct, at the common point, the *related helix** (that is the helix which agrees with the curve in osculating plane, curvature, and torsion), the axes of the helices so obtained generate a cylindroid.

13. The most general system with properties I and II involves two arbitrary functions of five arguments, besides two functions of three arguments. Two further statements, each equivalent to II, are as follows:

*An osculating helix, that is, one having four-point contact with the curve, does not in general exist.

If for each of the ∞^1 curves defined by a given lineal element we construct the osculating circle and the osculating sphere, the distance between the center of the circle and the center of the sphere varies as a linear integral function of the radius of curvature.

For the same set of ∞^1 curves, the derivative of the radius of curvature with respect to the arc length can be expressed as a linear integral function of the radius of curvature.

This last form has the advantage of being valid in space of two or any number of dimensions. On this basis, however, it would be difficult to formulate equivalents for the higher properties, so as to obtain a complete characterization.

14. Property III is perhaps the most interesting result obtained. The most general system having this property in addition to I and II is represented by a pair of differential equations involving ten arbitrary functions of three arguments.

One may ask what is the significance of the cubic curve (we denote it by Γ), which arises in connection with III. To each point O of space there is related a certain cubic Γ. If we shoot from O in every direction with every speed, we obtain ∞^3 trajectories. Each of these has an osculating sphere with a definite center C. To each of the trajectories there corresponds one center C. Usually the center C determines the trajectory. However if C lies on the curve Γ, there are ∞^1, instead of one, corresponding trajectory: in this case in fact the initial direction may be any direction perpendicular to the line joining O and C.* Thus the curve Γ may be defined as the locus of those points which may serve as centers for more than one trajectory through the given point O.

A simple consequence of III is that the locus of the centers of the osculating spheres of the ∞^2 trajectories touching a given plane at a given point is a quadric surface.

* Two trajectories through O have the same osculating sphere only if the initial speed is the same, and if the line through O perpendicular to the initial elements meets the cubic Γ.

If the plane varies, the given point being held fixed, the ∞^2 quadrics obtained form a linear system.*

The properties so far considered relate to the curves through a given point O. If we have ∞^3 curves passing through a point O, ∞^1 in each direction, and if, at that point, properties I, II, III are fulfilled, it will not usually be possible to generate the curves as trajectories in any field of force. All that follows is that the relations between y', z', y'', z'', y''', z''' are of precisely the same form as those holding for trajectories; and therefore it is possible to find (in infinitely many ways) a field of force such that each of the ∞^3 trajectories passing through the given point shall have contact of the third order with some one of the given curves.

15. In order to cause our system to be of the dynamical type, it is necessary to restrict the ten arbitrary functions involved in the type characterized by I, II, III so that only three arbitraries remain, namely, the components φ, ψ, χ defining the field of force. This is the rôle of property IV, which states that in any plane p the associated system S is of the plane dynamical type. An equivalent statement is as follows:

IV (2). If the ∞^2 space curves touching any plane p at any point O are projected orthogonally upon p, the plane curves thus obtained possess the properties I_p, II_p, III_p; when the point O varies in p, the direction associated with it by II_p, and the conic associated with it by III_p, vary in accordance with the restrictions expressed in IV_p and V_p.

It may be remarked that the first half of this statement holds for all space systems having properties I, II, III; in fact all such systems have also property IV_p. The real restriction is in V_p. It is also sufficient to consider, instead of all planes p, merely those of a triply orthogonal set.

§§ 16–25. The Inverse Problem of Dynamics: A Method of Geometric Exploration

16. The usual direct and inverse problems arising in dynamics are: first, given the force acting on a particle, to find its motion;

* On the other hand if we vary the given point, keeping the plane fixed, no simple result is obtained: the ∞^2 quadrics constitute an arbitrary family.

and second, given the motion of a particle, to find the force acting on it. The first problem is solved by integrating the differential equations of motion. The second is solved by differentiating the coordinates of the point with respect to the time.

Suppose, however, that we are given only the path described by the particle but have no information about the motion along the curve. If merely a single curve is given, the problem of finding the acting force would of course be indeterminate. But if all the trajectories, described by starting particles in a field of force from all initial positions in all directions with all speeds, are given, then the field of force is essentially determined (that is, up to a constant factor). *Hence if we were given a photograph of the entire system of curves generated by some (positional) field of force, without any record of motion or time, it ought to be possible to find the law of the field of force.* This is easily seen to be true analytically; but we wish also a purely geometric solution which will enable us to pass from the given curves to the vector representing the force at each point of the plane (taking first the two-dimensional case). The result gives what may be described as a method for the *geometric exploration of a field of force.*

17. First consider two trajectories passing through the same point O in the same direction. Construct the two osculating parabolas. The circle passing through the point O and the foci of these parabolas will, according to property I, be the focal circle corresponding to the given point and the given direction. Then, according to property II, the direction of the force acting at O will be symmetric to the tangent to this circle at O with respect to the common tangent of the two curves. An equivalent of this construction is to join O to the intersection of the directrices of the osculating parabolas: this line is perpendicular to the direction of the force acting at O.

If we are given two trajectories passing through O in different directions, then the direction of the force at O is not determined. The same is true if we are given three curves with distinct tangents.

18. *If, however, we are given four trajectories with distinct tangents, the force direction is (in general) uniquely determined.*

Consider an arbitrary direction at O, and let us see if it can be the direction of the force acting at that point. Take the image of this direction in the tangent to the first of the given curves; then pass a circle through O in the direction so obtained and through the focus of the corresponding osculating parabola. Doing this for each of the four curves, we obtain four focal circles. *If there exists a circle touching these four, the direction tested is correct.* This follows from property III (4) of § 8. We have then a purely geometric problem: to find a direction at O such that the four circles constructed by means of it shall admit a common tangent circle. We may simplify this problem by inverting the configuration considered with respect to O. We then have, instead of the four focal circles, four straight lines which are to be concyclic. As we change the direction tested, these rotate simultaneously through equal angles about four fixed points, namely, those obtained by inverting the four foci.

Take an arbitrary oriented direction for trial; construct for each of the four inverse foci, a direction parallel to the image of the tangent to the focal circle with respect to the tangent to the corresponding trajectory. We thus obtain four oriented lineal elements, one at each of the inverse foci. The problem is then to rotate these through the same angle α, so that the new elements shall have concyclic lines.* In this position the image of the direction of any one of the four elements in the corresponding tangent at O will give the required direction of the force. The only ambiguity, in general, will be in the sense (arrowhead) of the force: this, however, may be determined separately for actual† trajectories by considerations of concavity and convexity.

19. The direct analytical treatment is as follows. The differential equation of the ∞^3 trajectories of any positional field

* A simple ruler and compass solution of this problem was suggested to the author by Professor Wedderburn.

† See § 9.

of force is of the form

$$(y' - \omega)y''' = (\lambda y'^2 + \mu y' + \nu)y'' + 3y''^2,$$

where λ, μ, ν, ω are functions of x, y (and have therefore fixed numerical values so far as we deal with the ∞^2 curves passing through a given point O), the latter quantity ω representing the slope of the acting force. Each of the four given curves C_1, C_2, C_3, C_4 through the point O determines certain values of the derivatives y', y'', y'''; that is we are given the differential elements of third order

$$y_i', \quad y_i'', \quad y_i''' \qquad (i = 1, 2, 3, 4).$$

Substituting these values we have four linear equations

$$(y_i' - \omega)y_i''' = (\lambda y_i'^2 + \mu y_i' + \nu)y_i'' + 3y_i''^2 \quad (i = 1, 2, 3, 4),$$

from which we can find the values of λ, μ, ν, ω at the given point. *The required direction of the acting force is determined by the slope*

$$\omega = \frac{\begin{vmatrix} 1 & y_i' & y_i'^2 & \dfrac{y_i'y_i''' - 3y_i''^2}{y_i''} \end{vmatrix}}{\begin{vmatrix} 1 & y_i' & y_i'^2 & \dfrac{y_i'''}{y_i''} \end{vmatrix}},$$

where numerator and denominator are determinants of the fourth order.

20. By any of these methods we may determine the direction* of the vector representing the force acting at any point O of the plane. How shall we determine the magnitude of the vector? The determination cannot be absolute, since, as already remarked, two fields that differ by a constant factor have identical trajectories. The magnitude of the vector at any one point may be taken at random, and then the field is completely determined.

This depends upon the simple fact that if we know the *path*

* Of course if *all* the trajectories were given, the direction of the force would be determined immediately by the fact that the curves in that direction have zero curvature.

of a particle and also the *direction* of the force acting at each of its points, then assuming the *magnitude* arbitrarily at one point, it is completely determined at all points. This is an integration problem. We know the force vector at the initial point O, and may decompose it into components N and T, normal and tangent to the given curve. Assuming the mass to be unity, the initial speed is given by

$$v_0^2 = rN,$$

where r is the known radius of curvature. Then from

$$vv_s = T,$$

we may find v_s, the rate of variation of the speed for unit of arc. The speed at any point P of the curve is thus found in the form

$$v = v_0 e^{\int \frac{T}{rN} ds},$$

where all the quantities in the right-hand member are geometrically given. (The integrals throughout are calculated from point O to point P.) If we denote by θ the inclination of the force to the curve, so that $\tan\theta = N/T$, the speed is

$$v = v_0 e^{\int \frac{\cot\theta}{r} ds}.$$

Since the speed, that is the motion, is now known, the magnitude of the force is also known. The components at any point P are

$$N = N_0 e^{\int \frac{2\cot\theta - r_s}{r} ds}, \quad T = N \cot\theta.$$

21. We see that the construction of the field may be carried out without knowing all the ∞^3 trajectories. So far as the direction of the force is concerned, it is sufficient to know at each point of the plane either two trajectories with a common tangent, or four trajectories with distinct tangents. So far as

the magnitude is concerned it is then sufficient to know ∞^1 trajectories through one point O, one for each direction, since we can then integrate from this point to any point of the plane* along some one of the curves.

A field of force is in general determined, and may be constructed, if we know $4\infty^1$ out of the totality of ∞^3 trajectories, each of the four systems of ∞^1 curves covering the plane (or the region considered) simply, that is, one passing through each point of the plane.

The complete system of ∞^3 trajectories is thus determined in general by four systems of ∞^1 trajectories. Further reduction is possible. In general $3\infty^1$ curves determine the totality, but no simple constructions are then available. If two simply infinite systems of curves (that is, a net of curves) are assigned arbitrarily, a corresponding complete system can be found in a large infinitude of ways: the corresponding field of force is not determined up to a constant, but involves arbitrary functions.

The first and most interesting example of the geometric exploration of a field of force arose in Bertrand's discussion of Kepler's laws. The first of these laws (every planet describes an ellipse having the sun for a focus) is geometric, while the second and third are kinematic (involving the areal velocity and the period). The first law determines all the trajectories, and therefore determines the field of force.† Hence the newtonian law of gravitation can be deduced from the first law alone, instead of, as usual, from all three. Bertrand thus concludes that the other two laws are consequences of the first. If Kepler had been a mathematician of the twentieth century, he would have stopped his laborious observational inductions after noting his first law, and deduced the other two analytically.

The first law, in Bertrand's discussion, is of course to be taken ideally: not only the actual planets describe conics with a focus at the sun, but every particle starting from any position with

* That is, in some region of the plane—in some neighborhood of O.

† It is assumed, of course, that the force depends only upon the position of the planet.

any velocity describes such a conic. From what has been stated above it is sufficient to limit the observations to four simply infinite systems of conics in "general" position.

On account of the last phrase, it is easily possible to commit errors in the application of the result. It would be possible to give $4\infty^1$ or even ∞^2 conics in certain special ways, so that the field is *not* determined. (See § 23.)

22. This raises the general question: *How many trajectories may be common to two distinct fields of force?*

The first field, defined by its components φ, ψ, has a system of ∞^3 trajectories with a differential equation

$$(y' - \omega)y''' = (\lambda y'^2 + \mu y' + \nu)y'' + 3y''^2;$$

the second field, with components φ_1, ψ_1, has a system of trajectories given by an analogous equation

$$(y' - \omega_1)y''' = (\lambda_1 y'^2 + \mu_1 y' + \nu_1)y'' + 3y''^2.$$

If there are any solutions in common,* they must satisfy the equation of second order

$$3(\omega - \omega_1)y'' = (y' - \omega)(\lambda_1 y'^2 + \mu_1 y' + \nu_1) - (y' - \omega_1)(\lambda y'^2 + \mu y' + \nu).$$

Two systems of trajectories cannot have more than ∞^2 curves (one through each point in each direction) in common without coinciding. If they have ∞^2 curves in common the differential equation of the second order defining these curves must be of the cubic form†

$$y'' = Ay'^3 + By'^2 + Cy' + D,$$

where the coefficients are functions of x, y.

Usually the solutions of the equation of the second order will

* In addition to straight lines, $y'' = 0$, which are common to *all* systems.

† This form is characterized by the fact that the locus of the centers of curvature of the curves passing through a given point is a special type of cubic curve. Cf. *Amer. Jour. Math.*, 1908, p. 207.

not satisfy either equation of the third order and the two systems will have no curves in common. An example showing that the two systems may actually have ∞^2 curves in common is given by the fields

$$\varphi = x, \quad \psi = 4y; \qquad \varphi_1 = x^{-3}, \quad \psi_1 = 1,$$

where the equation of second order,

$$xy'' = y',$$

defines ∞^2 curves $y = ax^2 + b$, which are trajectories in both fields.

23. A fortiori $4\infty^1$ curves, or any number of simple systems, *may* belong to two distinct fields. If the four simple systems are given in the form

$$y' = f_i(x, y) \qquad (i = 1, 2, 3, 4),$$

the field, if it exists, will be *uniquely* defined provided not all the determinants of fourth order in the matrix

$$\| f_i', \quad f_i f_i', \quad f_i^2 f_i', \quad f_i'', \quad 3f_i'^2 - f_i f_i'' \|$$

vanish identically. Here the primes denote complete differentiation with respect to x, so that

$$f' = f_x + f f_y,$$
$$f'' = f_{xx} + 2f f_{xy} + f^2 f_{yy} + f_x f_y + f f_y^2.$$

This is the exact formulation of the result stated previously "in general."

24. Consider the simplest of all fields, gravity assumed constant. If a cannon ball is projected in any way into the field it describes a vertical parabola. Conversely if every path in an unknown field is a vertical parabola, it follows that the acting force is vertical and constant in intensity. *How many cannon ball experiments would have to be made in order to arrive at this conclusion?*

We confine the discussion for simplicity to a fixed vertical

plane, taken as the xy-plane, so that the equations of motion are

$$\ddot{x} = 0, \qquad \ddot{y} = 1$$

and the trajectories are the ∞^3 parabolas

$$y = ax^2 + bx + c.$$

Suppose first the cannon is kept in one place, say the origin, and the ball is fired in all directions with all initial speeds, giving in all ∞^2 parabolas

$$y = ax^2 + bx.$$

This would not be sufficient to prove that the field is uniform. Another possible field, for example, is

$$\ddot{x} = x^{-5}, \qquad \ddot{y} = yx^{-6}.$$

In fact there are ∞^2 distinct fields each consistent with the given set of ∞^2 parabolas.

The same is true if we confine our geometric experiments to the ∞^2 parabolas $y = ax^2 + c$ found by shooting horizontally from every point in the axis of ordinates with variable initial speed. The differential equation of this family is $xy'' = y'$, precisely the one given at the end of § 22, and so the two forces there given are consistent with the experiments, just as much as ordinary gravity.

If however the shots are fired from all points in the axis of abscissas, with all initial speeds, at the fixed inclination of 45°, producing as trajectories the ∞^2 vertical parabolas whose foci are on the axis of abscissas, *the field must be uniform gravity*. The only possible field is in fact $\ddot{x} = 0$, $\ddot{y} =$ constant.

The same is true if we fix the amount of powder, that is the initial speed, and shoot from every point on the ground (the axis of x), at every angle. This gives ∞^2 parabolas with a common directrix.

As an example of a set of $4\infty^1$ observations that would be sufficient, we mention only the case of shooting from four*

* It may be that three stations are sufficient, but this requires a separate discussion. Two stations would certainly not suffice.

stations on the ground, pointing the cannon at the angle 45°, and using all initial speeds.

25. Consider very briefly the general inverse problem in space of three dimensions. The determination of the magnitude of the force involves the same considerations as in the plane case.

If we are given two trajectories through O in the same direction, the osculating planes must coincide. The force acts in this common plane; its direction is determined by projecting the given space curves orthogonally on this plane, and then using the plane construction described above.

If we are given two trajectories with distinct osculating planes, the initial directions will be necessarily distinct; the force-direction is then determined by the intersection of the osculating planes.

If we are given two trajectories through O in different directions, but with the same osculating plane, the direction of the force is not determined. We need in fact four such curves with the same osculating plane and different directions before the force-direction is determined: the requisite construction is again obtained by orthogonal projections of the curves of the common osculating plane, thus reducing the problem to that considered in the two-dimensional theory (cf. § 18).

§§ 26 27. Tests for a Conservative Field

26. Since the system of trajectories determines the field of force, it ought to be possible to find out from the trajectories, whether the field belongs to any special type, for example, whether the field is central or conservative.

The lines of force are determined geometrically by property I in the plane and property II in space. The field will be central if the lines of force are straight lines passing through a common point.

We now give a number of tests any one of which will distinguish a conservative from a non-conservative force. It is not possible to decide this from the lines of force alone.

1°. First consider the *plane theory.* Here there is for each point a certain conic determined by the trajectories in accordance with property III of § 3 as the locus of the centers of the hyperosculating circles. *For a conservative field (and for no other) this conic is always a rectangular hyperbola.*

2°. In connection with property III (3) of § 8 we have this test: The conic which there appears as the locus of the centers of the focal circles is in the conservative case merely a straight line. That is, the focal circles constructed at any point all have a second point in common.

3°. The focal circles corresponding to two perpendicular directions are, in any field, tangent to each other. In the conservative case the two circles coincide.

4°. In any field two trajectories through a given point O exist whose osculating parabolas have the same given focus. If for one given focus the trajectories are orthogonal at O, this will be true for any given focus. When this is the case for every point O, the force will be conservative.

27. In the *three-dimensional theory*, the lines of force in the conservative case necessarily form a normal congruence; but this is not a sufficient test. All the tests given below are both necessary and sufficient.

1°. First consider property III of § 11. In any field there corresponds to each point O a certain twisted cubic curve Γ. The conservative fields are distinguished by the fact that the cubic Γ is, for every point O, of the rectangular type.*

2°. An interesting kinematic test, connected with the theorem of Thomson-Tait, is the following. If from any point O we shoot with a given speed v_0 in every direction, ∞^2 trajectories will be obtained. If these form a normal congruence (that is admit a set of orthogonal surfaces), the same will necessarily be true for any other speed v_0. *The trajectories starting out from any point with*

* That is, the cubic intersects the plane at infinity in three mutually orthogonal directions. All the quadrics passing through the curve are then of the equilateral type.

a given speed form a normal congruence when, and only when, the field is conservative.

The necessity of this condition is included in the Thomson-Tait theorem discussed in the next chapter. Its sufficiency, of course, requires a separate discussion which is connected with the theory of velocity systems.

3°. In order to make the preceding test purely geometric, it is necessary to have a geometric method of assembling those trajectories which, starting from the same point, correspond to the same initial speed. Such a method is readily found from the fact that the square of the speed varies directly as the radius of curvature and directly as the normal component of the force. The ∞^2 trajectories corresponding to a given speed have circles of curvature intersecting each other at the same point on the line of the force vector; that is, the centers of curvature lie in a plane perpendicular to the direction of the force acting at the given point. In the conservative case, the ∞^2 trajectories so selected form a normal congruence.

4°. Among the ∞^2 trajectories considered there are, for any field, three which admit hyperosculating circles of curvature. The three initial directions thus determined will be mutually orthogonal when and only when the field is conservative.

Only test 1° is directly connected with the set of properties I–IV of page 19. The other three are suggested by the discussion of velocity systems (cf. § 32).

CHAPTER II

NATURAL FAMILIES: THE GEOMETRY OF CONSERVATIVE FIELDS OF FORCE

§ 28. Origin and Application of the Natural Type

28. We now consider the properties of the trajectories generated by conservative fields of force. The total system of trajectories will have the general properties previously considered for an arbitrary field of force, together with the additional properties stated in §§ 26, 27, peculiar to the conservative case.

An entirely new feature presents itself, due to the fact that the differential equations of motion admit an integral of the first order, namely, the energy equation. During any motion of the particle in the given field, the sum of the kinetic and potential energies is constant; thus each motion corresponds to a definite value of the constant h, representing the total energy. The motions may therefore be grouped according to the values of h. Those corresponding to a given value form what may be termed, following Painlevé, a *natural* family.

Thus, in space of two dimensions, the complete system of trajectories for a given conservative field of force consists of ∞^3 curves grouped into ∞^1 natural families, each composed of ∞^2 curves. For example, in the case of ordinary gravity the trajectories are the ∞^3 vertical parabolas (in a given vertical plane), and the natural families are formed by grouping together those parabolas which have the same (horizontal) line as directrix.

In space of three dimensions, the complete system contains ∞^5 trajectories grouped into ∞^1 natural families, each containing ∞^4 curves. Examples are the ∞^4 parabolas with vertical axes whose directrices are situated in a fixed horizontal plane; and the ∞^4 circles orthogonal to a fixed sphere. The simplest example, corresponding to the case of zero force, is the ∞^4 straight lines of space.

This grouping of the trajectories according to the values of the total energy constant, that is, into natural families, is fundamental in most dynamical investigations relating to conservative forces, in particular, those connected with the principle of least action and the developments of Hamilton and Jacobi. From this point of view, dynamical problems relating to the same field of force, but having distinct values of h, are considered as essentially distinct problems. Quoting Darboux: "This restriction is in accordance with the spirit of modern mechanics which attaches less importance to force than to energy, and which permits us to regard as distinct two problems in which the force function or work function is the same, but the total energy is different."

It therefore seems of interest to work out the purely geometric properties of natural families. According to the principle of least action, such a family is made up of the extremals defined by the variation problem

$$\int \sqrt{W + h}\, ds = \text{minimum},$$

that is, the curves which cause the first variation of the integral to vanish. This follows from the fact that the speed v, in the action integral $\int v ds$, is determined by the energy equation

$$v^2 = 2(W + h).$$

Abstractly, a natural family of curves may be defined as one which can be regarded as the totality of extremals connected with a variation problem of the form

$$\int F ds = \text{minimum},$$

where F is any point function, that is, any function of x, y, z in the three-dimensional case.

Such families arise not only in the discussion of trajectories, but also, for example, in the discussion of brachistochrones, catenaries, optical rays, geodesics, and contact transformations.

The brachistochrone problem for a conservative field with any work function W leads to the integral

$$\int dt = \int \frac{ds}{\sqrt{W + h}}.$$

Thus the complete system of brachistochrones is made up of ∞^1 natural families, one for each value of h.

When a homogeneous, flexible, inextensible string is suspended in the conservative field, the forms of equilibrium, which are termed catenaries in the general sense of the word, are obtained by rendering the integral

$$\int (W + h) ds$$

a minimum. Hence here also we have ∞^1 natural families, one for each value of h.*

Consider an isotropic medium in which the index of refraction ν varies arbitrarily from point to point. The paths of light in such a medium, according to Fermat's principle of least time, are determined by minimizing the integral $\int \nu ds$ and hence form a single natural family. This is the most concrete way of defining a natural family.

The connection with the theory of geodesics is obvious. Thus in the two-dimensional case the geodesics of the surface whose squared element of length (first fundamental form) is $\lambda(x, y)(dx^2 + dy^2)$ are found by minimizing the integral $\int \sqrt{\lambda}\, ds$, and hence the representing curves in the x, y plane constitute a natural family. Hence if any surface is represented conformally on a plane, the geodesics are pictured by a natural family of curves in that plane. The extension to more variables is evident:

* The complete systems of ∞^5 brachistochrones and ∞^5 catenaries have geometric properties distinct from each other and from those of the ∞^5 trajectories: no quintuply infinite system of curves can be at the same time the system of trajectories for some field and the system of brachistochrones or catenaries in either the same or a different field. The distinctive properties for an arbitrary field are given in § 107, p. 94. Cf. § 103.

any natural family in any space may be obtained by conformal representation from the geodesics of some other space.*

As a last application we consider the transformations which Sophus Lie has termed the infinitesimal contact transformations of mechanics. In the plane case, such a transformation is defined by a characteristic function of the special form $\Omega(x, y)(1 + y'^2)^{\frac{1}{2}}$ and is characterized by the fact that the lineal elements at each point are converted into the elements of a circle about that point as center. The path curves of every contact transformation of this category form a natural family.

§§ 29–31. Characteristic Properties A and B

29. *Osculating Circles—Property A.*—We now consider the general geometric properties of a natural family in ordinary space, that is, the totality of ∞^4 extremals connected with an integral of the form

$$\int F(x, y, z)\sqrt{1 + y'^2 + z'^2}\, dx. \tag{1}$$

The differential equations of the family are then the corresponding Euler-Langrange equations

$$\begin{aligned} y'' &= (L_y - y'L_x)(1 + y'^2 + z'^2), \\ z'' &= (L_z - z'L_x)(1 + y'^2 + z'^2), \end{aligned} \tag{2}$$

where

$$L = \log F.$$

Of the ∞^4 curves in this family ∞^2 pass through any given point p, one in each direction. Our first result is:

Theorem 1: *The ∞^4 curves in any natural family have this property: the circles which at any point p of space osculate the ∞^2 curves passing through that point, have a second point P in common and thus form a bundle.*

* A natural family on a given surface may be regarded as a family of pseudo-geodesics, that is, one which may be obtained as the conformal picture of the geodesics on some other surface.

This property we shall refer to as *property A*. In the discussion it is convenient to decompose it into these two statements, also relating to the ∞^2 curves through a given point:

(A_1) The osculating planes constructed at the common point form a pencil.

(A_2) The centers of curvature lie in a plane perpendicular to the axis of the pencil of osculating planes.

A proof of the theorem stated is easily obtained by regarding the family as made up of dynamical trajectories. Property A_1 results from the fact that the osculating plane of a trajectory always passes through the force vector. Property A_2 is proved by noting that those trajectories through a given point, which correspond to the same value of the total energy h, are all described with the same initial velocity v_0. The radius of curvature at the initial point is given by the formula

$$r = v_0^2/N,$$

where N denotes the component of the force along the principal normal. Since N is the orthogonal projection of a fixed vector, the locus of its terminal point will be a sphere through the initial point. The conclusion then follows from the fact that r varies inversely as N.

The following analytical discussion has the advantage of answering the converse question which naturally arises: Are there other systems with property A?

The differential equations of any system of ∞^4 space curves, one determined by each lineal element of space, may be assumed in the form

$$(3)\qquad y'' = f(x, y, z, y', z'), \qquad z'' = g(x, y, z, y', z').$$

Property A_1 requires that at each point there shall be a certain direction through which all the osculating planes at that point must pass. Let the direction in question be given by the ratios of three arbitrary point functions

$$(4)\qquad \phi(x, y, z), \qquad \psi(x, y, z), \qquad \chi(x, y, z);$$

then the requisite condition is

$$\frac{z''}{y''} = \frac{\chi - z'\phi}{\psi - y'\phi}. \tag{5}$$

Property A_2 requires that the centers of curvature shall lie in a plane perpendicular to the direction (4); hence

$$\phi X + \psi Y + \chi Z = 1, \tag{6}$$

where X, Y, Z denote the coordinates of the center relative to axes with the common point as origin. Using the general formulas for the center of curvature, and combining with (5), we find

THEOREM 2: *The differential equations of any system of curves possessing property A are of the form*

$$\begin{aligned} y'' &= (\psi - y'\phi)(1 + y'^2 + z'^2), \\ z'' &= (\chi - z'\phi)(1 + y'^2 + z'^2), \end{aligned} \tag{7}$$

where ϕ, ψ, χ are arbitrary functions of x, y, z. The converse is valid also.

The equations (2) are seen to be included in this form, hence the result certainly holds for our natural systems, as stated in theorem 1.

30. *Hyperosculations—Property B.*—The circles of curvature at a given point, for any system of the form (7), constitute a bundle. We now inquire whether any of these circles correspond to four-point, instead of three-point, contact.

If a twisted curve is to have an hyperosculating circle of curvature at a given point, two conditions must be satisfied, namely,

$$\begin{vmatrix} 1 & y' & z' \\ 0 & y'' & z'' \\ 0 & y''' & z''' \end{vmatrix} = 0, \tag{8}$$

$$\frac{dr}{ds} = 0. \tag{9}$$

The first of these states that the osculating plane has four-point contact with the curve; the second, in which r denotes the radius of curvature, is the condition for the existence of an osculating helix, i. e., one with four-point contact. When both conditions hold the helix is simply the circle of curvature, which then has hypercontact.

Applying these conditions to the curves defined by (7), we find, from (8),

(10) $$(\psi - y'\phi)\chi' - (\chi - z'\phi)\psi' + (y'\chi - z'\psi)\phi' = 0;$$

and, from (9),

(11) $$(1 + y'^2 + z'^2)\Sigma\phi\phi' - (\phi + y'\psi + z'\chi) \times \left\{ \begin{array}{c} (1 + y'^2 + z'^2)\Sigma\phi^2 + \phi' + y'\psi' + z'\chi' \\ - (\phi + y'\psi + z'\chi)^2 \end{array} \right\} = 0,$$

where the indicated summations extend over ϕ, ψ, χ and where ϕ', for example, denotes $\phi_x + y'\phi_y + z'\phi_z$.

Since we wish to discuss the ∞^2 curves through a given point, we may simplify our equations considerably by taking the axis of abscissas in the special direction (4). Then, at the selected point, ψ and χ vanish, and the above equations reduce to

(10′) $$y'\chi' - z'\psi' = 0,$$

(11′) $$(y'^2 + z'^2)(\phi' - \phi^2) - (y'\psi' + z'\chi') = 0.$$

Neglecting the trivial solutions for which $y'^2 + z'^2$ vanishes, we may reduce this pair of simultaneous equations to the form

(12) $$\frac{\psi_x + y'\psi_y + z'\psi_z}{y'} = \frac{\chi_x + y'\chi_y + z'\chi_z}{z'} = \phi_x + y'\phi_y + z'\phi_z - \phi^2.$$

This set of equations for the determination of y', z' is of a familiar type, namely, that arising in the determination of the fixed points of a collineation, and is easily shown to admit three solutions.* Hence

* Of course in special cases some of these may coincide, or the number of solutions may become infinite. The theorem stated is true "in general" in so far as it omits these cases which are definitely assignable.

THEOREM 3: *The curves defined by equations of the form* (7) *are such that through each point there pass three with hyperosculating circles at that point.*

Since the form (7) is characterized by property A, it follows that the existence of three hyperosculating circles in each bundle is a consequence of property A.

We state two further properties, found by considering the conditions (10′) and (11′) separately.

The tangents to those curves of a system (7) *which pass through a given point and there have an hyperosculating plane form a quadric cone.* This cone passes through the special direction (4).

The tangents to those curves which have an osculating helix at the given point form a cubic cone. This cone passes through the special direction (4) and through the minimal directions in the plane normal to that direction.

These properties hold for natural families since they hold for all systems with property A. By comparing (7) with (2), we see that the functions ϕ, ψ, χ in the case of a natural family are

$$(13) \qquad \phi = L_x, \qquad \psi = L_y, \qquad \chi = L_z;$$

and hence are connected by the relations

$$(14) \qquad \psi_z - \chi_y = 0, \qquad \chi_x - \phi_z = 0, \qquad \phi_y - \psi_x = 0.$$

We now inquire what is the effect of these relations on the directions of the hyperosculating circles. Introducing, for symmetry,

$$(15) \qquad X : Y : Z = 1 : y' : z',$$

we may write our equations (12) in the homogeneous form

$$(16) \qquad \begin{aligned} &\chi_y Y^2 - \psi_z Z^2 + (\chi_z - \psi_y) YZ + \chi_x XY - \psi_x XZ = 0, \\ -&\chi_x X^2 + \phi_z Z^2 + \phi_y YZ - \chi_y XY + (\phi_x - \phi^2 - \chi_z) XZ = 0, \\ &\psi_x X^2 - \phi_y Y^2 - \phi_z YZ - (\phi_x - \phi^2 - \psi_y) XY + \psi_z XZ = 0. \end{aligned}$$

In virtue of (14), each of the quadric cones (16) is seen* to be

* The condition for such a cone is that the sum of the coefficients of X^2, Y^2, and Z^2 shall vanish.

of the rectangular type. Hence the three generators common to the cones must be mutually orthogonal. This gives

THEOREM 4: *In the case of any natural family the three hyperosculating circles which exist in any bundle are mutually orthogonal.*

We refer to this property as *property B*.

31. The relations (14) are seen to be necessary as well as sufficient for the orthogonality in question. Hence property B is the equivalent of (14), and serves to single out the natural families from the more general class defined by equations of form (7). The latter form was characterized by property A; hence we have our

FUNDAMENTAL THEOREM: *A system of ∞^4 curves, one for each direction at each point of space, will constitute a natural family when, and only when, it possesses properties A and B: that is, the osculating circles at any given point must form a bundle, and the three hyperosculating circles contained in such a bundle must be mutually orthogonal.*

§ 32. General Velocity Systems

32. The most general system with property A is represented by differential equations of the form

$$(7) \qquad \begin{aligned} y'' &= (\psi - y'\phi)(1 + y'^2 + z'^2), \\ z'' &= (\chi - z'\phi)(1 + y'^2 + z'^2), \end{aligned}$$

and thus involves three arbitrary functions. Only in the case where these functions are the partial derivatives of the same function is the system a natural one. We now point out a dynamical problem that leads to the general type (7): this justifies the term *velocity system* which we hereafter employ to denote any system of this type.

Consider a particle (of unit mass) moving in any field of force, the components of the force being ϕ, ψ, χ. The equations of motion are then

$$\ddot{x} = \phi(x, y, z), \qquad \ddot{y} = \psi(x, y, z), \qquad \ddot{z} = \chi(x, y, z).$$

If the initial position and the initial velocity are given the motion

is determined. If only the initial position and direction of motion are given, the osculating plane will be determined but the radius of curvature r will depend for its value on the initial speed v. Hence, in addition to the usual formula

$$v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2,$$

there must be a formula expressing v^2 in terms of x, y, z, y', z', r. This is furnished by the familiar equation

$$v^2 = rN,$$

where N denotes the (principal) normal component of the force, so that

$$N^2 = \phi^2 + \psi^2 + \chi^2 - \frac{(\phi + y'\psi + z'\chi)^2}{1 + y'^2 + z'^2}.$$

The result may be written in the two (equivalent) forms

$$v^2 = \frac{(\psi - y'\phi)(1 + y'^2 + z'^2)}{y''} = \frac{(\chi - z'\phi)(1 + y'^2 + z'^2)}{z''}.$$

In the actual trajectory v varies from point to point. If now we replace v^2 in this result by some constant, say $1/c$, the resulting equations may be written

$$y'' = c(\chi - y'\phi)(1 + y'^2 + z'^2),$$
$$z'' = c(\chi - z'\phi)(1 + y'^2 + z'^2).$$

The curves satisfying these differential equations—they are not in general trajectories—we define as *velocity curves*. For any field a curve is a velocity curve corresponding to the speed v_0, provided a particle starting from any lineal element of the curve with that speed describes a trajectory osculating the curve. In a given field of force there are ∞^5 trajectories and ∞^5 velocity curves.* If c is given we have ∞^4 velocity curves. In particular

* The properties of a complete system of ∞^5 velocity curves are analogous to, but distinct from, those of a complete system of trajectories. Cf. p. 94.

if c (and hence v) is taken to be unity, our equations become precisely (7).

Any system of ∞^4 curves possessing property A, that is, any system (7), may be regarded as the totality of velocity curves corresponding to unit velocity in some (uniquely defined) field of force.

Only when the field is conservative do the velocity systems for each value of v (or c) become natural systems. The trajectories also are in this case made up of ∞^1 natural families, one for each value of the energy constant h; but the two sets of natural families are distinct. The determination of a velocity system in one conservative field is equivalent to the determination of a trajectory system in another conservative field, and vice versa. We find in fact the following explicit result:

*If two conservative fields with work functions W_1 and W_2 satisfy the relation**

$$W_2 = ae^{\frac{2W_1}{v_0^2}} - h,$$

then the ∞^4 velocity curves for the speed v_0 in the first field coincide with the ∞^4 trajectories for the constant of energy h in the second field.†

§ 33. Reciprocal Systems

33. With any velocity system S

$$(S) \qquad y'' = (\psi - y'\varphi)(1 + y'^2 + z'^2), \quad z'' = (\chi - z'\varphi)(1 + y'^2 + z'^2)$$

there is connected a definite point transformation T: for in virtue of property A to any point p corresponds a definite point P, the osculating circles constructed at the first point all passing through the second point. The transformation T is explicitly

$$(T) \qquad X = x + \frac{2\varphi}{\varphi^2 + \psi^2 + \chi^2}, \cdots .$$

* We note that if W_1 is left unaltered and v_0 varied, W_2 takes quite distinct forms. The ∞^1 velocity systems in a given field do not constitute the complete system of ∞^5 trajectories in any field whatever.

† It is seen that the two fields have the same equipotential surfaces and therefore the same lines of force. (Central fields therefore correspond to central fields.)

It is thus entirely general. To an arbitrary transformation* corresponds a definite velocity system. In particular, to the inverse transformation T^{-1} there corresponds a certain system S', which we define as reciprocal to S.

Hence to a general† velocity system S, that is, any system possessing property A, there corresponds a definite reciprocal velocity system S'. The osculating circles of those curves of system S which pass through any point p are at the corresponding point P the osculating circles of the curves of the system S' passing through P.

Consider the bundle of circles determined by two corresponding points p and P. We know that three of these circles have hypercontact with S-curves at p, and three have hypercontact with S'-curves at P. It is not obvious that the circles so obtained really coincide. Omitting the rather long proof, we merely state the result.

Reciprocal velocity systems have the same hyperosculating circles: the three circles hyperosculating curves of the given system S at any point p also hyperosculate curves of the reciprocal system S' at the corresponding point P.

It follows at once that if S possesses property B (that is mutually orthogonal hyperosculating circles) the same will be true of S'. This means that whenever system S is natural so is S'.

The reciprocal of a natural family is always a natural family.

We may restate this in optical terms as follows: With any isotropic medium, defined by its index of refraction $\nu(x, y, z)$, there is connected a certain *reciprocal medium* with an index of refraction $\tilde{\nu}(x, y, z)$: the rays of light in this second medium, namely, the extremals of

$$\int \tilde{\nu}(x, y, z)ds = \text{minimum},$$

form the system reciprocal to that formed by the rays of light

* It may even degenerate but must not be merely the identical transformation. We however exclude systems with degenerate T's from the rest of the discussion: we assume that the jacobian does not vanish, so that the inverse transformation exists.

† See preceding footnote.

in the given medium, namely, the extremals of

$$\int \nu(x, y, z) ds = \text{minimum}.$$

The actual calculation of ν from v requires only operations that are *performable* in the Lie sense, namely, eliminations and differentiations. See *Transactions of the American Mathematical Society*, volume 10 (1909), page 213.

§ 34. Character of the Transformation T

34. The transformation T (from point p to point P) associated with the most general system possessing property A is, as we have seen, entirely arbitrary. The question arises what is the peculiarity of T if the given system is of the natural type. The answer to this will furnish an equivalent of property B, and will thus make it possible to characterize natural families without introducing hyperosculating circles.

The problem is to describe geometrically the class of transformations of the form

$$X = x + \frac{2L_x}{L_x^2 + L_y^2 + L_z^2}, \quad Y = y + \frac{2L_y}{L_x^2 + L_y^2 + L_z^2},$$
$$Z = z + \frac{2L_z}{L_x^2 + L_y^2 + L_z^2},$$

depending on one arbitrary function L of x, y, z, instead of three independent functions required in a general point transformation,

$$X = \Phi(x, y, z), \qquad Y = \Psi(x, y, z), \qquad Z = \mathrm{X}(x, y, z).$$

For a general (analytic) point transformation the bundle of lineal elements at any point is converted linearly into the bundle at the corresponding point. Are there any elements which go over into parallel elements? It is well known that there are three. If in particular these three elements are mutually perpendicular (for every point of space), we obtain a certain category of transformations which may be termed Darboux* transforma-

* See *Proceedings of the London Mathematical Society*, 1900.

tions or deformations. They are analytically of the form

$$X = f_x, \qquad Y = f_y, \qquad Z = f_z,$$

involving one arbitrary function. Obviously this is not the class we desire.

We next ask whether in the general transformation there are any elements at a given point p each of which is turned into a *cocircular* element at the corresponding point P. This is, in a way, a case correlative to the Darboux case: for whether two elements in space are parallel or cocircular they have in common the properties that they are coplanar and equally inclined to the line pP joining their points. It is found that there are always three such elements at any point. If we require these to be mutually orthogonal, we obtain precisely the transformations connected with natural families.

A system of ∞^4 space curves possessing property A will form a natural family when and only when the associated transformation T (from point p to point P) has the following property: the three lineal elements at p each of which is converted into a cocircular element at P are mutually orthogonal.

We have thus obtained an equivalent for property B. It may be shown synthetically that the three directions just described (cocircular elements) always coincide with the directions of the hyperosculating circles. The orthogonality of the one triple amounts to the same thing as the orthogonality of the other.

It may be remarked that the class of transformations connected with all natural systems do not form a group. It is obvious however that the inverse of any member of the class is contained in that class. *This is the essence of the law of reciprocity for natural systems, previously obtained by a different method.*

§§ 35–44. The Converse of Thomson and Tait's Theorem

35. It is well known that if straight lines are drawn orthogonal to any given surface they will necessarily be orthogonal to an

infinitude of surfaces (namely the surfaces parallel to the given surface). Thomson and Tait in their Natural Philosophy showed that this property of the ∞^4 straight lines of space holds for the ∞^4 trajectories described in any conservative field with the same total energy, that is, for any natural family. The writer has proved that no other families of curves have the property: it is entirely characteristic of the natural type.* We first state the original theorem in connection with the general theory of the calculus of variations, and then take up the converse theorem. Later a second converse question is discussed.

35′. *Thomson and Tait's Theorem.* We have seen that a natural family of curves in space may be regarded as the totality of extremals of a variation problem of the particular form

$$J = \int F(x, y, z) ds, \tag{1}$$

where F is a point function, ds is the element of length

$$ds = \sqrt{dx^2 + dy^2 + dz^2} = \sqrt{1 + y'^2 + z'^2}\, dx,$$

and the integral is taken between fixed end points.

It is easily shown that for integrals of this form,† and for no others, the relation of *transversality*, in the sense of the calculus of variations, amounts merely to *orthogonality*. This suffices to distinguish our type among variation problems of the general form

$$\int f(x, y, z, y' z') dx. \tag{2}$$

But of course it does not serve as a complete geometric test for a natural family. What is the geometric character of the systems of ∞^4 extremals connected with any variation problem (2)? This is an unsolved question in the calculus of variations.‡

* At least in the case of space of three dimensions. Cf. *Trans. Amer. Math. Soc.*, vol. 11 (1910), pp. 121–140.

† Cf. Bolza, Variationsrechnung, p. 691; also p. 146 for the two-dimensional problem due to Hedrick.

‡ See the author's paper, "Systems of extremals in the calculus of variations," *Bull. Amer. Math. Soc.*, vol. 13 (1908), pp. 289–292.

We are concerned here only with the integrals of special form J, defining natural families. Applying Kneser's fundamental theorem on transversals,* we have this well-known result: If from the points of any surface Σ we construct the extremals orthogonal to the surface, and on each lay off an arc so that the integral J takes some constant value, then the locus of the end points is a surface which is also orthogonal to the extremals.

36. This is known as the *theorem of Thomson and Tait.* It was obtained by them in connection with the dynamics of a particle moving in a conservative field—the first interpretation of a natural family considered in § 28. Here $F(x, y, z)$ represents the speed v, as determined by the energy equation

$$v^2 = 2(W + h),$$

where W denotes the work function (negative potential), and the mass is assumed to be unity. Of course h has a fixed value. We quote the original statement of the theorem:

"If from all points of an arbitrary surface particles not mutually influencing one another be projected normally with the proper velocities [so as to make the sum of the kinetic and potential energies have a given value]; particles which they reach with equal actions lie on a surface cutting the paths at right angles."

The integral J, in this case, represents the action

$$\int v\,ds = \int \sqrt{2(W + h)}\,ds.$$

The ∞^1 surfaces cutting the curves orthogonally thus appear as surfaces of equal action.

The corresponding statement for brachistochrones is sometimes called the *theorem of Bertrand*:† From the points of any surface draw the brachistochrones normal to the surface and on each lay off lengths so that the time of transit is equal to a given quantity; then the locus of the end points will be another surface orthogonal

* Bolza, pp. 131 and 691.

† Cf. Routh, Dynamics of a Particle (1898), p. 376. According to Appell, Mécanique rationelle, vol. 1 (1909), p. 466, this result was indicated by Euler.

to the brachistochrones. Here the integral J represents the time

$$\int dt = \int \frac{ds}{v} = \int \frac{ds}{\sqrt{2(W+h)}},$$

so that the orthogonal surfaces appear as surfaces of equal time.

Corresponding statements may be made, of course, for the other interpretations leading to natural families. The most concrete aspect is obtained by using the language of optics. Here the integrand function is simply the index of refraction $\nu(x, y, z)$, varying from point to point in any (isotropic) medium, and the integral $\int \nu ds$ is proportional to the time. The paths of light in such a medium form a (single) natural family, and every natural family may be obtained in this way. The ∞^2 rays (in general curved) starting out normally from anys urface admit ∞^1 orthogonal surfaces. These present themselves as surfaces of equal time. We shall describe them as a *set of wave fronts or wave surfaces.*

37. *The geometric part of the theorem of Thomson and Tait may be stated as follows: In any natural family of ∞^4 space curves, the ∞^2 curves which meet any surface orthogonally always form a normal congruence.*

Is this geometric property, which we shall refer to as the *Thomson-Tait property*, characteristic? This is in fact the case. We shall prove, namely, the following

CONVERSE THEOREM. *If a quadruply infinite system of curves in space is such that ∞^2 of the curves meet an arbitrarily given surface orthogonally* and always form a normal congruence (that is, admit an infinitude of orthogonal surfaces), then the system is of the natural type, that is, it may be identified with the extremal system belonging to an integral of the form $\int F(x, y, z)ds$.*

38. The result is simple but the proof is rather long. We give the essential steps.

Consider an arbitrary quadruply infinite system of curves in

* This means the same as requiring that one curve of the system passes through each point of space in each direction.

space, assuming that one passes through each point in each direction. Such a system may be defined by a pair of differential equations of the second order

(1) $$y'' = F(x, y, z, y', z'), \qquad z'' = G(x, y, z, y', z'),$$

where F and G are uniform functions which we assume to be analytic in the five arguments. Denoting the initial values of x, y, z, y', z', which may be taken at random, by x, y, z, p, q respectively, and employing X, Y, Z as current coordinates, we may write the solutions of (1) in the form

(2) $$\begin{aligned} Y &= y + p(X - x) + \tfrac{1}{2}F(X - x)^2 + \tfrac{1}{6}M(X - x)^3 + \cdots, \\ Z &= z + q(X - x) + \tfrac{1}{2}G(X - x)^2 + \tfrac{1}{6}N(X - x)^3 + \cdots. \end{aligned}$$

Here F and G are expressed as functions of x, y, z, p, q; and M and N, found by differentiating (1), are given by

(3) $$\begin{aligned} M &\equiv F_x + pF_y + qF_z + FF_p + GF_q, \\ N &\equiv G_x + pG_y + qG_z + FG_p + GG_q. \end{aligned}$$

The terms of higher order will not be needed in our discussion. Equations (2) involve five arbitrary parameters but of course represent only ∞^4 curves.

Consider now an arbitrary surface Σ

(4) $$z = f(x, y).$$

At each point of this surface and normal to it a definite curve of the given family (1) may be constructed. A certain congruence will thus be determined. We wish to express the condition that this shall be of the normal type, that is, that the ∞^2 curves shall admit a family of orthogonal surfaces.

The direction normal to the surface Σ at any point is given by

$$1 : p : q = f_x : f_y : -1,$$

so that

(5) $$p = P(x, y), \qquad q = Q(x, y),$$

where

(5′) $$P \equiv f_y/f_x, \qquad Q \equiv -1/f_x.$$

These functions are connected by the relation

$$(5'') \qquad PQ_x - QP_x - Q_y = 0.$$

The equations of the ∞^2 curves corresponding to the given initial conditions may now be written

$$(6) \qquad \begin{aligned} X &= x + t, \\ Y &= y + Pt + \tfrac{1}{2}\bar{F}t^2 + \tfrac{1}{6}\bar{M}t^3 + \cdots, \\ Z &= f + Qt + \tfrac{1}{2}\bar{G}t^2 + \tfrac{1}{6}\bar{N}t^3 + \cdots, \end{aligned}$$

where t takes the place of $X - x$ in (2), and where the bars indicate that the substitution (4), (5) has been carried out, so that, for example,

$$(7) \qquad \bar{F}(x, y) = F(x, y, f, P, Q).$$

The coefficients of the powers of t in (6) are thus functions of the two parameters x, y.

The general condition for a normal congruence given in parametric form is*

$$(8) \qquad (Y'XY) - (Z'ZX) + Y'(Z'YZ) - Z'(Y'YZ) = 0,$$

where the parentheses denote jacobians taken with respect to t, x, y, and Y', Z' denote the derivatives of Y, Z respectively with respect to t.

Expanding (8) in powers of t in the form

$$(9) \qquad \Omega_0 + \Omega_1 t + \Omega_2 t^2 + \cdots,$$

we find that Ω_0 vanishes in consequence of (5″). This is as it should be, since our ∞^2 curves are orthogonal to Σ by construction.

The terms containing the first power of t give

$$(10) \qquad \begin{aligned} &(1 + P^2 + Q^2)(P\bar{G}_x - Q\bar{F}_x - \bar{G}_y) \\ &\quad + 2\bar{F}\{(P^2+Q^2)(Q_x - PQ_y) + 2\bar{G}fPP_y - (P^2+Q^2)P_x\} = 0. \end{aligned}$$

* We may also use the convenient form due to Beltrami. Cf. Bianchi-Lukat, Differentialgeometrie, p. 340.

From (6′) we find

$$\overline{F}_x = F_x + F_z f_x + F_p P_x + F_q Q_x,$$

with corresponding results for $\overline{G}_x$ and $\overline{G}_y$. Substituting these values, and observing from (5) and (5′) that

$$f_x = -1/Q, \qquad f_y = -P/Q, \qquad Q_y = PQ_x - QP_x,$$

we may reduce (10) to

$$\begin{aligned} &2F\{PQP_x+Q^2Q_x\}+2G\{PP_y-(P^2+Q^2)P_x\}-(1+P^2+Q^2) \\ (10') \qquad &\times\{QF_x-F_z-PG_x+G_y+(QF_p-QG_q-PG_p)P_x \\ &\qquad\qquad +G_pP_y+QF_qQ_x\}=0. \end{aligned}$$

This is then a necessary condition in order that the ∞^2 curves belonging to the quadruply infinite system (1) and orthogonal to the surface (4) shall form a normal congruence. The result is to hold in virtue of (4) and (5).

It is of course not a sufficient condition. It merely expresses the fact that the curves orthogonal to Σ are also orthogonal to some consecutive surface, that is, that the congruence is approximately normal to the first degree.

Our main problem is to find all systems (1) which have the orthogonality property with respect to *every* base surface Σ. It is then necessary that (10′) should be true for an arbitrary function $f(x, y)$. The function can be so selected that for any chosen values of x and y the quantities f, P, Q, P_x, P_y, Q_x, shall take arbitrary numerical values; for the only relation to be fulfilled is (5″) and this merely determines Q_y. The condition (10′) must therefore hold identically. Arranging it in the form

$$(10'') \quad (1 + P^2 + Q^2)C_0 + QC_1Q_x + C_2P_y - C_3P_x = 0,$$

and equating coefficients to zero, we find

$$\begin{aligned} &C_0 \equiv qF_x - F_z - pG_x + G_y = 0, \\ &C_1 \equiv (1 + p^2 + q^2)F_q - 2qF = 0, \\ (11) \quad &C_2 \equiv (1 + p^2 + q^2)G_p - 2pG = 0, \\ &C_3 \equiv (1 + p^2 + q^2)\{qG_q + pG_p - qF_p\} \\ &\qquad\qquad + 2pqF - 2(p^2 + q^2)G = 0. \end{aligned}$$

Integration of the second and third of these partial differential equations gives

$$F = f_1(p, x, y, z)(1 + p^2 + q^2), \qquad G = g_1(q, x, y, z)(1 + p^2 + q^2),$$

where f_1 and g_1 denote unknown functions of the four arguments indicated. Substituting these values in the fourth equation, we find $f_{1p} = g_{1q}$, and therefore

$$f_1 = \psi - p\phi, \qquad g_1 = \chi - q\phi,$$

where ϕ, ψ, χ are functions of x, y, z only. The general solution of the last three equations of the set (11) is therefore

$$(12) \quad F = (\psi - p\phi)(1 + p^2 + q^2), \qquad G = (\chi - q\phi)(1 + p^2 + q^2).$$

We have still to satisfy the first equation of (11), which now reduces to

$$(13) \qquad \psi_z - \chi_y + p(\chi_x - \phi_z) + q(\phi_y - \psi_x) = 0.$$

The functions ϕ, ψ, χ must therefore satisfy the equations

$$(13') \quad \psi_z - \chi_y = 0, \qquad \chi_x - \phi_z = 0, \qquad \phi_y - \psi_x = 0,$$

and hence are expressible as the derivatives of a single function in the form

$$(13'') \qquad \phi = L_x, \qquad \psi = L_y, \qquad \chi = L_z.$$

The solutions of the set (11) are therefore

$$(14) \qquad \begin{aligned} F &= (L_y - pL_z)(1 + p^2 + q^2), \\ G &= (L_z - pL_x)(1 + p^2 + q^2), \end{aligned}$$

involving an arbitrary function L of x, y, z. The resulting system (1) is thus recognized to be a natural family. This gives our fundamental converse theorem.

39. In the above discussion use has been made, not of the complete condition for a normal congruence, but only of condition (10′) derived from the terms of the first order in t. We may therefore state a stronger converse result as follows:

The only systems of ∞^4 curves which have the property that the curves orthogonal to any surface are always orthogonal to some infinitesimally adjacent surface are those of the natural type.

If a congruence of curves meets two neighboring surfaces orthogonally it need not meet ∞^1 surfaces orthogonally, and therefore it approximates to, but need not coincide with, a normal congruence. The above theorem shows however that if the weak requirement of approximate normal character be imposed on *all* the congruences obtained from the given quadruply infinite system, they will *all* be exactly normal.

40. We may further strengthen our theorem by demanding the orthogonality property for some instead of all surfaces. Our fundamental equations (11) resulted from the fact that x, y, z, f, P, Q, P_x, P_y, Q_x might receive arbitrary numerical values. It will therefore be sufficient to take a manifold of surfaces sufficiently large to leave these quantities, or the equivalent quantities

$$(15) \qquad x, \quad y, \quad z, \quad z_x, \quad z_y, \quad z_{xx}, \quad z_{xy}, \quad z_{yy},$$

unrestricted. Since these quantities define a differential surface element of the second order, we may state the result as follows:

The converse theorem remains valid if, instead of considering all base surfaces, we employ a manifold of surfaces sufficiently large to include all the ∞^8 possible differential elements of the second order.

41. The Thomson-Tait theorem holds of course even when the base Σ shrinks to a curve or a point: there will still be a normal congruence orthogonal to the curve or point (in the latter case orthogonality means simply passage through the point). We state a number of results obtained in this connection.

If for an arbitrary curve as base the corresponding ∞^2 orthogonal curves of a given quadruply infinite system always form a normal congruence, the given system is necessarily natural.

If we require each of the congruences here considered to be of approximately normal character, a more general type of system

is obtained, namely the *velocity type* of § 32. The velocity type is thus characterized by the fact that those curves of the system which meet an arbitrary curve orthogonally are orthogonal to some infinitesimally adjacent (of course tubular) surface. We may even restrict ourselves to the case where the base is a curve of the given system, or the case where it is any straight line.

42. Suppose next that the base is an arbitrary point. Are natural families the only families of ∞^4 curves such that the ∞^2 curves passing through any point form a normal congruence? A discussion shows that this is not the case. There exist families not of the natural type, for example, that defined by the differential equations

$$y'' = y'^2, \qquad z'' = 0,$$

with the restricted property stated. To find all such systems would be a rather difficult, but certainly an interesting, undertaking. The result would of course include the natural type as a special case.

43. It will not however be the velocity type. It may be shown in fact that the only velocity systems for which the curves passing through an arbitrary point constitute always a normal congruence are those of the natural type. Recalling the fact that the velocity type is characterized by property A, we may give a new characterization of the natural type as follows:

Natural families are the only quadruply infinite systems of curves in space such that the ∞^2 curves through an arbitrary point admit an infinitude of orthogonal surfaces, and such that the osculating circles constructed at the common point form a bundle.

44. It may also be shown that if for every point and every straight line as base the corresponding congruence is normal, the system will be natural. To have a velocity system it is sufficient to demand that the congruence corresponding to an arbitrary straight line shall be approximately normal. To have a natural system it is sufficient to demand approximate normality for the congruences corresponding to arbitrary straight lines and planes.

§§ 45–53. Wave Propagation in an Isotropic Medium: Properties of Wave Sets

45. The optical interpretation of a natural family and the Thomson-Tait property suggest certain sets of surfaces which we shall now study.

Consider a given medium defined by its index of refraction $\nu(x, y, z)$ given as a function of position. The rays (in general curved lines) are the ∞^4 extremals of

$$(1) \qquad \int \nu(x, y, z)ds = \text{minimum};$$

they form the natural family, whose differential equations are

$$(2) \qquad \begin{aligned} y'' &= (L_y - y'L_x)(1 + y'^2 + z'^2), \\ z'' &= (L_z - z'L_x)(1 + y'^2 + z'^2), \end{aligned}$$

where

$$(2') \qquad L \equiv \log \nu.$$

The ∞^2 rays orthogonal to any selected surface Σ form a normal congruence, that is, are orthogonal to a set of ∞^1 surfaces. A disturbance originating in the medium on the surface Σ will be propagated in the medium through this set of surfaces, which we term a *set of wave fronts*. In the given medium an arbitrary surface belongs to one and only one of these wave sets. A single surface is thus of arbitrary character, but the sets of surfaces

$$(3) \qquad f(x, y, z) = \text{constant}$$

that may be wave sets are restricted by the Hamilton-Jacobi equation

$$(4) \qquad f_x^2 + f_y^2 + f_z^2 = \nu^2.$$

The given medium defines also a certain set of *level surfaces*

$$\nu(x, y, z) = \text{constant}.$$

This, it should be noticed, is not usually a wave set—the only exception arising when the level surfaces are parallel. For a given

medium the number of wave sets is ∞^∞, since there is one for each surface. Each of these sets is cut by the level surfaces in the equidistant curves of the wave set; that is, along any one of these curves the distance between consecutive wave surfaces remains the same.*

46. A single set of wave fronts has no geometric peculiarity. That is, given any set of surfaces $f(x, y, z) =$ constant, it will always be possible to find a medium in which that set will serve as a wave set. In fact there are ∞^∞ such media. For in equation (4), the given function f, without altering the given surfaces, may be replaced by an arbitrary function $\Omega(f)$ of itself, and this gives ∞^∞ distinct values for ν.

When will two sets of wave fronts be consistent? Two arbitrary sets of surfaces $f =$ constant, $f_1 =$ constant cannot usually be regarded as wave sets in any single medium. The requisite condition is

$$\frac{f_x^2 + f_y^2 + f_z^2}{f_{1x}^2 + f_{1y}^2 + f_{1z}^2} = \frac{\Omega(f)}{\Omega_1(f_1)},$$

where Ω, Ω_1 may be any functions. An equivalent condition is that it must be possible to chose parameters for the two sets in such a way that

$$\frac{df}{dn} = \frac{df_1}{dn_1},$$

where dn and dn_1 denote the normal distance between consecutive surfaces.

47. But a clearer answer may be given in terms of the geometric properties A and B. If a set of surfaces is to be a wave set, the ∞^2 orthogonal curves must be members of the natural family of ∞^4 rays. If two sets of wave fronts are given, we have then two congruences of curves. The question then is, when can two normal congruences of curves be regarded as belonging to a natural family?

* This follows immediately from (4). It is to be remarked, however, that this property is not characteristic of wave sets.

Take any point p in space, and consider the two curves, one from each of the congruences, passing through it. The circles of curvature at p must intersect again at some point P (by property A). This condition makes sure that the two congruences belong to some velocity system. If now this is to be a natural system, we must also add property B or rather, since no hyperosculating circles are directly defined, the equivalent restriction (see page 47) relating to the transformation from p to P. The final answer may then be given as follows:

Two sets of wave surfaces belong to the same optical medium when and only when they satisfy the following geometric conditions:

(A') *At any point p of space the circles of curvature of the orthogonal trajectories of the two sets of surfaces, passing through that point, intersect again at some point P.*

(B') *The point transformation from p to P has the property that the three lineal elements of p each of which corresponds to a cocircular element at P are mutually orthogonal.*

48. Two sets of surfaces taken at random will not belong, as wave sets, to any medium. On the other hand, as we have said, one set belongs to ∞^∞ distinct media. The question then arises, just what will uniquely determine a medium.

A natural family is uniquely determined if we are given one set of wave fronts and a single extra trajectory. This means a trajectory not belonging to the congruence defined as the orthogonal trajectories of the wave set.

49. The *extra curve* however cannot be taken at random; it must be related in a certain way to the wave set. If the wave set is $f(x, y, z) =$ constant, then the condition on the curve is that it satisfy the Monge equation of second order

$$(3)\qquad \begin{vmatrix} 2\Delta y'' - (1 + y'^2 + z'^2)(\Delta_y - y'\Delta_x) & f_y - y'f_x \\ 2\Delta z'' - (1 + y'^2 + z'^2)(\Delta_z - z'\Delta_x) & f_z - z'f_x \end{vmatrix} = 0,$$

where

$$(3')\qquad \Delta = f_x^2 + f_y^2 + f_z^2.$$

Here f, and hence Δ, are given, and y and z are unknown functions of x. The interpretation is obvious from property A.

In order that an extra curve shall be consistent with a given wave set (that is, in order that both shall belong to a single medium) it is necessary and sufficient that the curve shall cross the surfaces (of course obliquely) in such a way that at any point of intersection the circle of curvature of the extra curve shall intersect the circle of curvature of the curve orthogonal to the surfaces. When the curve satisfies this restriction, it defines with the given wave set a unique natural family.

50. If we are merely given one wave set, the number of possible media is ∞^{∞} (since ν involves arbitrary functions). Each of these has ∞^4 rays (forming a natural family). The totality of media give rise to a totality of ∞^{6} rays, namely the solutions of the Monge equation of second order (3). This equation is of the type

$$Ay'' + Bz'' + C = 0$$

(where the coefficients are functions of x, y, z, y', z'), which the author has shown to be characterized by the *Meusnier property:** Those curves which pass through a given point in a given direction have circles of curvature (constructed at the common point) generating a sphere.

51. The inverse problem connected with natural families, namely, given the ∞^4 trajectories to construct the generating field of force, is solved immediately in connection with property A. The force acting at any point p acts in the line joining that point to the corresponding point P, and its intensity is proportional to the reciprocal of the distance between the two points.† This construction may be carried out if we know a sufficient number of trajectories, without knowing the whole system.

52. The greatest number of rays which two distinct media

* Kasner, *Bull. Amer. Math. Soc.*, vol. 14 (1908), pp. 461–465. The result includes the extension of Meusnier's theorem made by Lie, and is in fact the largest generalization possible.

† The determination of the potential function $W(x, y, z)$ or, what is equivalent, the index of refraction $\nu(x, y, z)$, requires a quadrature.

can have in common is ∞^2 (one through each point of space). If two media have that many in common, it is easily shown that the resulting congruence is necessarily normal. Any normal congruence can be obtained in this way, for, as stated above, it belongs, not only to two, but to ∞^∞ distinct media.

53. We mention only one special problem: the determination of those media in which disturbances are propagated by Lamé families of surfaces; that is, every wave set is to be of the Lamé type (thus forming part of a triply orthogonal family of surfaces). The index of refraction is found to vary inversely as the *power* of the point with respect to a fixed sphere; the rays then are the ∞^4 circles orthogonal to that sphere. Since the radius of the sphere may be zero, real, or imaginary, these media yield well known interpretations of parabolic, hyperbolic, and elliptic geometries. (See *Transactions of the American Mathematical Society*, volume 12 (1911), pages 70–74.)

§§ 54–61. A Second Converse Problem Connected with the Thomson-Tait Theorem

54. Consider the general conservative field, defined by its work function $W(x, y, z)$. With any motion of the particle there is associated a definite value of the constant of total energy

$$\tfrac{1}{2}v^2 - W = h.$$

If h is not assigned the complete system of trajectories is made up of ∞^5 curves.

Consider now an arbitrary surface, which we term the base surface,

$$(\Sigma) \qquad z = f(x, y).$$

From each of its points we may draw normal to the surface ∞^1 trajectories since the initial value of the speed v is arbitrary. We thus have in all ∞^3 trajectories normal to Σ. In order to have a congruence we must assign the value of v at each point of Σ, that is, we must give a law of distribution of the initial speed. The question arises: What form of law will make the corresponding

congruence a normal congruence? Of course for any law the congruence will be orthogonal to the base surface, but usually it admits no other orthogonal surfaces.

The Thomson-Tait theorem (in its complete dynamical form) gives *one* such law: it states that if the initial speed is selected so as to make h have the same value at all the points of Σ, the congruence will be normal. It thus gives a plan for constructing ∞^1 normal congruences for a given base, one for each value of h. We shall refer to any one of these as "constructed according to the Thomson-Tait law."

Is this the only answer to our question? If ∞^2 trajectories are drawn orthogonal to Σ and if they form a normal congruence, does it follow that the distribution of values of the initial speed is precisely such that the sum of the kinetic and potential energies has the same value at all points of Σ?

The requisite discussion is not simple. We shall merely state the results we have obtained.

55. The answer to our question is "in general" in the affirmative. The first converse theorem, discussed in § 37, is true without exception. The present is true with exceptions which may be definitely limited.

For a "general" base surface Σ in a given conservative field of force, the only congruences, formed by ∞^2 trajectories orthogonal to Σ (one drawn at each point), which admit an infinitude of orthogonal surfaces, are those constructed according to the Thomson-Tait law (so that the total energy has a constant value).

56. To make this precise we must of course limit the class of *exceptional surfaces* connected with a given field. These appear in the analytical discussion as the solutions of a certain partial differential equation of the second order*

$$\begin{vmatrix} W_x + pW_y & \overline{W}_x + p\overline{W}_y \\ W_x + qW_z & \overline{W}_x + q\overline{W}_z \end{vmatrix} = 0,$$

* The expanded result is of the form

$$P_1r + P_2s + P_3t + P_4 = 0,$$

where r, s, t denote the derivatives of second order of $z = f(x, y)$.

where W is the given work function, and

$$\overline{W} = \frac{pW_x + qW_y - W_z}{\sqrt{1 + p^2 + q^2}}.$$

This differential equation defines a class of surfaces which is seen to depend only on the equipotential surfaces

$$W(x, y, z) = \text{constant}.$$

The result may be put into geometric form and stated as follows:

The only surfaces Σ which may be exceptional in the theorem of § 55 (that is, which may give rise to normal congruences not included in the Thomson-Tait law) are those with this property: along each of the equipotential lines of the surface the component of the acting force normal to the surface is constant.*

57. Observe that it is not stated that the surfaces described, which exist in any field, actually give rise to additional normal congruences. To understand the situation more precisely, it is necessary to observe that in the analytic discussion the condition for a normal congruence is developed in the form

$$t\,\Omega_1 + t^2\Omega_2 + \cdots = 0,$$

where t is the parameter which varies along the curve, starting with the value zero on the surface Σ, and the coefficients Ω are functions of the two parameters defining the initial points on Σ. By assumption the congruence is orthogonal to Σ, so the term Ω_0, independent of t, will not appear. For a normal congruence all the coefficients Ω must vanish. If only a certain number vanish the congruence may be described as *approximately normal* (the approximation being of degree n if $\Omega_1 = \Omega_2 \cdots = \Omega_n = 0$): the curves are then orthogonal not only to Σ but also to one or more (infinitesimally) adjacent surfaces.

58. If now we impose on the congruence of trajectories normal to Σ the condition $\Omega_1 = 0$, we find that this may be fulfilled for

* The equipotential lines of any surface are the lines cut out by the equipotential surfaces W = const.

any surface: the restriction is merely on the law of initial speed and means that the total energy must be the same, not necessarily over the entire surface, but along each equipotential line of the surface.*

59. If we further impose the condition $\Omega_2 = 0$, then for a "general surface" the law of speed must be the Thomson-Tait law, but for an "exceptional surface" the law is the more general one just stated.

60. The discussion of the higher conditions $\Omega_3 = 0$, etc., we have not completed. It is therefore not known precisely in which cases normal congruences (in the exact sense) may arise. For central and parallel fields it may be shown that the exceptional surfaces† actually give rise to normal congruences (in addition to those included in the Thomson-Tait theory): for such fields the vanishing of the higher coefficients follows from the vanishing of the first two.

61. The principal results of the converse problem may be formulated as follows:

If ∞^2 trajectories (of a conservative field), meeting a surface Σ orthogonally, are also orthogonal to an infinitesimally adjacent surface, then the total energy along each equipotential line of Σ is constant.

If ∞^2 trajectories, selected from the complete system of ∞^5, form a normal congruence, then in general they will all belong to the same natural family (that is, the total energy will be the same for all the curves); except possibly when the ∞^1 orthogonal surfaces‡ are exceptional in the sense defined in § 56 (the additional congruences then and only then are normal to at least the second degree of approximation).

Normal congruences not of the Thomson-Tait type (that is, not

* If, in particular, the surface is one of the equipotential surfaces, the distribution of speed is thus entirely arbitrary.

† In the case of ordinary constant gravity the exceptional surfaces are those termed *moulure* surfaces by Monge: they are generated by rolling the plane of any plane curve about a vertical cylinder of arbitrary cross section.

‡ If one of these surfaces is exceptional, all will be.

selected from within a natural family) actually arise for central and parallel fields.

§§ 62–67. Geometric Formulation of Some Curious Optical Properties

62. In Thomson and Tait's Natural Philosophy* the characteristic function of Hamilton is applied to the motion of a particle in a conservative field of force, and certain results are obtained which we shall try to restate as purely geometric properties of a natural family of trajectories. To what extent these properties are characteristic is not settled. We quote the principal passages referred to.

"Let two stations, O and O', be chosen. Let a shot be fired with a stated velocity, V, from O, in such a direction as to pass through O'. There may clearly be more than one natural path by which this may be done; but, generally speaking, when one such path is chosen, no other, not considerably diverging from it, can be found; and any infinitely small deviation in the line of fire from O, will cause the bullet to pass infinitely near to, but not through, O'. Now let a circle, with infinitely small radius r, be described round O as center, in a plane perpendicular to the line of fire from this point, and let—all with infinitely nearly the same velocity, but fulfilling the condition that the sum of the potential and kinetic energies is the same as that of the shot from O—bullets be fired from all points of this circle, all directed infinitely nearly parallel to the line of fire from O, but each precisely so as to pass through O'. Let a target be held at an infinitely small distance, a', beyond O', in a plane perpendicular to the line of the shot reaching it from O. The bullets fired from the circumference of the circle round O, will, after passing through O', strike this target in the circumference of an exceedingly small ellipse, each with a velocity (corresponding of course to its position, under the law of energy) differing infinitely little from V', the common velocity with which they pass through O'. Let now a circle, equal to the former, be described round O',

* Part I (Cambridge, 1903), pp. 355–359.

in the plane perpendicular to the central path through O', and let bullets be fired from points in its circumference, each with the proper velocity, and in such a direction infinitely nearly parallel to the central path as to make it pass through O. These bullets, if a target is held to receive them perpendicularly at a distance $a = a'V'/V'$, beyond O, will strike it along the circumference of an ellipse equal to the former and placed in a "corresponding" position; and the points struck by the individual bullets will correspond; according to the following law of "correspondence":—Let P and P' be points of the first and second circles, and Q and Q' the points of the first and second targets which bullets from them strike; then if P' be in a plane containing the central path through O' and the position which Q would take if its ellipse were made circular by a pure strain; Q and Q' are similarly situated on the two ellipses."

63. The second passage is as follows: "The most obvious optical application of this remarkable result is, that in the use of any optical apparatus whatever, if the eye and the object be interchanged without altering the position of the instrument, the magnifying power is unaltered." . . . "Let the points O and O' be the optic centers of the eyes of two persons looking at one another through any set of lenses, prisms, or transparent media arranged in any way between them. If their pupils are of equal size in reality, they will be seen as similar ellipses of equal apparent dimensions by the two observers. Here the imagined particles of light, projected from the circumference of the pupil of either eye, are substituted for the projectiles from the circumference of either circle, and the retina of the other eye takes the place of the target receiving them, in the general kinetic statement."*

* This fact and many other applications are included in the following general proposition. "The rate of increase of any one component momentum, corresponding to any one of the coordinates, per unit of increase of any other coordinate, is equal to the rate of increase of the component momentum corresponding to the latter per unit increase or dimension of the former coordinate, according as the two coordinates chosen belong to one configuration of the system, or one of them belongs to the initial configuration and the other to the final."

64. The statement in the first passage is not purely geometric; for it involves not only the curves described, but also the speeds V and V' at the points O and O'. We therefore try to formulate the part of the theorem which is really geometric.

We have a natural family made up of ∞^4 curves in space, one for each initial lineal element (point and direction) of space. Select any one of these curves c and any two points O and O' upon it. Construct the planes p and p' normal to this curve at O and O'.

For each direction through O, a curve of our family is determined; this strikes the plane p' at a definite point. We thus have a certain correspondence between the bundle of directions through O and the points of p'. For directions infinitesimally close to the direction of c at O, and for points close to O', this correspondence is linear; and by a proper selection of cartesian axes at O and O', we may write the correspondence in the canonical form

$$\xi = \alpha_1 x', \quad \eta = \beta_1 y',$$

where (x', y') denote the coordinates of the point in the plane p', and the corresponding direction at O has direction cosines proportional to $(\xi : \eta : 1)$.

In an entirely analogous way, by considering the curves of the natural family which go through O', and the points of intersection with the plane p, we obtain a second linear correspondence which may be reduced to the form

$$\xi' = \alpha_2 x, \quad \eta' = \beta_2 y,$$

where (x, y) is the point in the plane p and $(\xi' : \eta' : 1)$ gives the corresponding direction at O'.

If we were dealing with an arbitrary family of ∞^4 curves, instead of a natural family, these linear correspondences would still exist; but the choice of axes in the second canonical form would be different from that required in the first, and the two constants appearing in the second form would be independent of those

appearing in the first. *The peculiarity of the natural type may be stated in the following form: First, the canonical axes for the two correspondences coincide; second, the ratio of the characteristic constants has the same value for both correspondences.*

This is the essential geometric content of the long statement quoted above from Thomson and Tait. Is this characteristic of the natural type? We do not know.

64′. A statement in more concrete terms is of interest. If we start out from O in directions equally inclined (the fixed angle is of course assumed infinitesimal) to the direction of c, that is, along a cone of revolution having for axis the tangent of c, the resulting trajectories forming a sort of curvilinear cone, we strike points on p' located on an ellipse with O' as center. By changing the angle of the cone we obtain a family of similar and similarly situated ellipses. The principal axes of these ellipses are the canonical directions referred to above for the first correspondence, and the ratio of the diameters is equal to the ratio of the canonical constants $(\alpha_1 : \beta_1)$. By starting from the other point O' along cones of revolution having for axis the tangent to c, we strike the plane p in a second set of homothetic ellipses. *The two sets of ellipses thus obtained, one in the plane p, and the other in the plane p', are similar.* This is part of the property stated, but not the whole. It should be observed that it has no meaning to say that the two sets are similarly situated, since they are in different planes.

65. We may, however, obtain two sets in the same plane as follows: If we start along the cone of revolution from O, we hit p' in an ellipse. If we wish to hit p in a circle, we must start at O' along a certain elliptical cone: the sections of this cone by planes parallel to p', projected orthogonally on p', give a set of homothetic ellipses. We thus have in the plane p', two sets of ellipses, the first set being obtained from cones of revolution at O, and the second set being obtained from elliptical cones at O' by orthogonal projection of parallel sections. If we were dealing with an arbitrary family of curves, the two sets thus obtained would be unrelated: *for a natural family, however, the two sets coincide.*

66. Of course we could also construct two sets in the plane p and these would coincide; but this would not give an additional property. In the statement quoted, certain pairs of congruent instead of merely similar ellipses appear, but that is due to the introduction of kinematics: namely, use is made of the velocities V and V' at the points O and O'. "If O and O' are regarded as optic centers of the eyes of two persons looking at one another through any optical apparatus, and if their pupils are of equal size in reality, they will be seen as similar ellipses of equal apparent dimensions by the two observers." It should be observed, however, that the dimensions will be *equal* only under the assumption that the two eyes are at positions for which the velocities V and V', or, what is equivalent, the indices of refraction ν and ν', are equal. In the most general case of an isotropic medium, the ellipses will not have equal apparent dimensions, but the ratio of the dimensions will be equal to the ratio of the two velocities.

67. Two converse questions remain unanswered. First: Find all systems of ∞^4 curves in space such that circles about O and O' appear as similar ellipses.

Second: Find all systems such that the set of ellipses in the plane p' formed by starting from O along cones of revolution, and the set of ellipses found by orthogonal projection upon p' of the sections cut out by planes parallel to p' of those (elliptical curvilinear) cones at O' which strike plane p in circles,—such that these two sets of ellipses shall coincide.

§§ 68-72. The So-called General Problem of Dynamics

68. Consider any material system (particles or rigid bodies) with n degrees of freedom, so that its position at each instant is determined by n independent coordinates denoted by $x_1, x_2, \cdots, x_n$. The kinetic energy T will be represented by a quadratic form

$$2T = \sum a_{ik}\dot{x}_i\dot{x}_k,$$

where the coefficients a are functions of the coordinates, and

the dots denote time derivatives. If the acting forces are conservative, there will exist a force function $W(x_1, x_2, \cdots, x_n)$, which is assumed to be independent of the time, and the equation of energy

$$T - W = h$$

asserts that in any given motion the sum of the kinetic and potential energies is constant.

The so-called general problem of dynamics requires the determination of the motions when we are given the form T, the function W, and the constant h. The possible trajectories are then given by the Jacobi principle of least action as the extremals of the integral

$$\int \sqrt{W + h}\, \sqrt{\Sigma a_{ik} dx_i dx_k}.$$

This defines the *most general natural family*. The integral is of the form $\int F ds$, where F is any point function and ds is the length-element in a general n-dimensional variety V_n defined by

$$ds^2 = \Sigma a_{ik} dx_i dx_k.$$

69. Such a family consists of $\infty^{2(n-1)}$ curves, in the space V_n, one passing through each point in each direction. A *complete characterization* is given by J. Lipke, in his doctor's dissertation,* as follows:

(A_1) The locus of the centers of geodesic curvature of the ∞^{n-1} curves passing through any point of V_n is a flat space of $n - 1$ dimensions S_{n-1}.

(A_2) The osculating geodesic surfaces (two-dimensional varieties) at the given point form a bundle of surfaces, all containing a fixed direction (and hence the geodesic line in that direction) which is normal to the S_{n-1} of property A_1.

(B) The n directions at any point, in which, as a consequence of the preceding properties, the osculating geodesic circles (circles

* *Trans. Amer. Math. Soc.*, vol. 13 (1912), pp. 77–95.

of constant geodesic curvature) hyperosculate the curves of the given family, are mutually orthogonal.

70. This gives the generalization of properties A and B stated in §§ 29–31. The simpler results there given for ordinary space apply to a euclidean space of any dimensionality and also to spaces of constant curvature. In the general space of variable curvature, the geodesic circles constructed at a given point do not all meet at a second point, and so no analogue of the law of reciprocity of natural families presents itself.

71. The theorem of Thomson and Tait remains valid for any space.* The converse questions connected with it have not been settled. In all probability the Thomson-Tait geometric property is characteristic in any space (flat or curved) of dimensionality greater than two. Obviously in the case of two dimensions the geometric converse is not valid, since any system of ∞^1 curves admits ∞^1 orthogonal curves.

72. The systems characterized by property A (meaning A_1 together with A_2) are the most general velocity systems in V_n. The case $n = 2$ presents a peculiar feature: for then, included in the velocity type, we have, in addition to the natural type, another special type of interest (geometric, rather than dynamic), namely the isogonal type† (systems formed by the ∞^2 isogonal trajectories of an arbitrary simply infinite system of curves). In the case of the plane (or any surface of constant curvature) the reciprocity construction for velocity systems is available, and each of the species, natural and isogonal, is self-reciprocal. The only families common to the two species are those formed by the isogonals of an isothermal system, or, what is the same, by velocity systems generated by Laplacian fields of force.‡

* Cf. Darboux, *Leçons*, vol. 2, last chapter, where references to the memoirs of Lipschitz and Beltrami are given.

† Scheffers introduced the systems of plane curves $y'' = (\psi - y'\varphi)(1 + y'^2)$ in connection with the theory of isogonals, and obtained a law of reciprocity for isogonal systems. Cf. *Leipziger Berichte*, 1898, 1900; *Mathematische Annalen*, vol. 60.

‡ Cf. the author's note, "Isothermal systems in dynamics," *Bull. Amer. Math. Soc.*, vol. 14 (1908), pp. 169–172.

We note finally this characteristic distinction between the two noteworthy species:

For both natural and isogonal families in the plane, the circles of curvature constructed at any point p have another point P in common. The point transformation T (from p to P) in the natural case is such that the two lineal elements at any point, each of which is converted into a cocircular element, are orthogonal; while in the isogonal case the two elements, each of which is converted into an element normal to a cocircular element, are orthogonal.

If the transformation T connected with a velocity system is required to be (direct) conformal, the corresponding field must be Laplacian. Such fields are distinguished from all others by the fact that each of the infinitude of systems of velocity curves is then expressible linearly in the two parameters involved.

CHAPTER III

TRANSFORMATION THEORIES IN DYNAMICS

§§ 73–81. Projective Transformations

73. The general object of a transformation theory is to relate new problems to old problems, and so to proceed from the solution of the latter to the solution of the former. The most important geometric transformations are the projective and the conformal. Both groups play important rôles in dynamics, the former in connection with general fields, and the latter in connection with conservative fields.

74. The importance of projective transformations in dynamics was brought out by Appell in 1889. Given any positional field of force in the plane, the corresponding equations of motion are of the form

$$(1) \qquad \frac{d^2x}{dt^2} = \varphi(x, y), \qquad \frac{d^2y}{dt^2} = \psi(x, y).$$

If an arbitrary point transformation, unaccompanied by any change in the time, is applied, the new differential equations will usually involve not only x and y, but also the velocity components dx/dt, dy/dt. In fact the only exception is where the point transformation is merely affine:

$$x_1 = ax + by + c, \quad y_1 = a'x + b'y + c'.$$

Appell showed that if a general collineation

$$(2) \qquad x_1 = \frac{ax + by + c}{a''x + b''y + c''}, \qquad y_1 = \frac{a'x + b'y + c'}{a''x + b''y + c''}$$

is accompanied by a change of the time of the form

$$(2') \qquad dt_1 = \frac{dt}{k(a''x + b''y + c'')^2},$$

the new differential equations will be of the original form

$$(3) \qquad \frac{d^2x_1}{dt_1^2} = \varphi_1(x_1, y_1), \quad \frac{d^2y_1}{dt_1^2} = \psi_1(x_1, y_1),$$

and therefore define motion in some new positional field of force. The relation between the new field and the original field is explicitly as follows

$$(4) \qquad \begin{aligned} \varphi_1 &= k^2(a''x + b''y + c'')^3\{C''(x\psi - y\varphi) + B'\varphi - A'\psi\}, \\ \psi_1 &= k^2(a''x + b''y + c'')^3\{-C'(x\psi - y\varphi) - B\varphi + A\psi\}, \end{aligned}$$

where the capital letters denote minors in the determinant $|ab'c''|$ of (2).

74′. The trajectories of the original field are converted by the collineation into the trajectories of the new field. Also the directions of forces of the two fields are projectively related. It must not be thought, however, that the force vector acting at a given point (x, y) in the first plane is projected into the new force vector acting at the point (x_1, y_1) of the second plane: the initial points of the two vectors will correspond, of course, by the given collineation, but the terminal points will not. The question therefore arises, *what is the geometric relation between the new vector field and the old vector field?*

To answer this question we take our rectangular axes so that the collineation takes its metrical normal form. (Affinities of course require a separate discussion.) The canonical formulas for our transformation are

$$(5) \qquad \begin{gathered} x_1 = \frac{\gamma\gamma_1}{x}, \quad y_1 = \frac{\gamma_1 y}{x}, \\ \varphi_1 = -k^2\gamma\gamma_1 x^3\varphi, \quad \psi_1 = k^2\gamma_1 x^3(x\psi - y\varphi), \end{gathered}$$

together with

$$(5') \qquad dt_1 = \frac{dt}{kx^2}.$$

To each collineation between the two planes corresponds a defi-

nite vector transformation. The vectors are here of the third type (*bound* vectors) described in the Introduction, requiring four coordinates for their determination. The original vector is defined by the four numbers (x, y, φ, ψ), the first two defining the initial point, and the last two giving the components of the vector. The coordinates of the new vector are $(x_1, y_1, \varphi_1, \psi_1)$.

The vector transformation induced by the given collineation is not projective. The new vector has the same initial point and the same direction as the projection of the old vector, but has a different length. The ratio λ between the actual length of the new vector and the length of the projected vector is

$$\lambda = k^2x^3(x+\varphi). \tag{5''}$$

Noting that in the canonical form x and $x+\varphi$ denote the distances from the initial and terminal points of the original vector to the vanishing line in the first plane, we may state this result.

Any given (non-affine) collineation (2) induces a certain vector transformation (determined up to the factor k) defined analytically by (2) and (4), and geometrically as follows: If PQ is any bound vector in the first plane, and if the collineation converts the initial point P into P_1 and the terminal point Q into Q_1, then the transformed bound vector is not P_1Q_1, but P_1Q_1', where Q_1' is the point on the line joining P_1Q_1 such that the ratio $\lambda = P_1Q_1'/P_1Q_1$ equals k^2 times the cube of the distance from P to the vanishing line times the distance from Q to that vanishing line.*

The transformation converts the ∞^4 bound vectors of the first plane, represented by the independent coordinates (x, y, φ, ψ), into the ∞^4 bound vectors of the new plane.† In the dynamical application, φ and ψ are given as functions of x, y, that is, we have a field of ∞^2 vectors, one for each initial point: the result of

* In the case of an affine collineation, the induced vector transformation is, except for the constant factor k, merely the result of applying the affinity to both ends of the vector. It is thus linear.

† The vector transformations induced by inverse collineations are inverse to each other. The four-dimensional transformations are therefore Cremona transformations.

the transformation is a new field, φ_1 and ψ_1 being expressible in terms of x_1, y_1. The ∞^3 trajectories of the first field are converted by the collineations into the ∞^3 trajectories of the new field; it is to be noticed however that, during any corresponding motions, positions which correspond according to the collineation will usually not correspond to the same instant of time; in fact from (2′)

$$t_1 = \int \frac{dt}{k^2(a''x + b''y + c'')^2}.$$

75. If X, Y denote the velocity components at the position x, y and if the corresponding velocity in the second plane is X_1, Y_1, acting at the position x_1, y_1, then we find, from the canonical form (5),

$$(6) \qquad \begin{aligned} x_1 &= \gamma\gamma_1/x, & y_1 &= \gamma_1 y/x, \\ X_1 &= -k\gamma\gamma_1 X, & Y_1 &= k\gamma_1(xY - yX). \end{aligned}$$

Thus we have a different vector transformation which may be termed the *phase* transformation* (in distinction from the *force transformation* of § 74): it gives the relation between the corresponding phases in the two planes.

If we speak of points and vectors which correspond in the two planes according to the given collineation as projectively related, then the result may be stated in this form:

The new phase vector does not coincide with the projection of the given phase vector: it has the same initial point, but the ratio of the actual length to the length of the projected vector is k^2 times the product of the distances from the ends of the original vector to the vanishing line of the collineation.

76. Having studied the Appell transformation and its geometric interpretation in terms of force vectors and phase vectors, we now ask whether other more general transformations can play a like rôle. Appell proved the following converse theorem:

* The phase of a particle at any instant, in the sense of Gibbs, is its position together with its velocity: it is defined by the four numbers $(x, y, \dot{x}, \dot{y})$.

The only transformations of the form

$$x_1 = \Phi(x, y), \qquad y_1 = \Psi(x, y), \qquad dt_1 = \mu(x, y)dt$$

which convert every set of differential equations

$$(1) \qquad \frac{d^2x}{dt^2} = \varphi(x, y), \qquad \frac{d^2y}{dt^2} = \psi(x,y),$$

into one of the same form are those defined by (2), (2′).

77. By eliminating the time from (1), giving the differential equation of the trajectories in the form (page 7)

$$(7) \quad (\psi - y'\varphi)y''' = \{\psi_x + (\psi_y - \varphi_x)y' - \varphi_y y'^2\}y'' - 3\varphi y''^2,$$

the author proved that the only point transformations which convert every trajectory system (of a positional field) into a trajectory system are the collineations. This remains valid even in the domain of all contact transformations, as we now proceed to show.

We first consider the class of differential equations (cf. page 11)

$$(8) \qquad y''' = G(x, y, y')y'' + H(x, y, y')y''^2$$

including (7) as a special case, and characterized geometrically by the possession of property I (that is, the focal locus for each element is a circle through the given point). We prove this theorem:

The only contact transformations which convert every equation of type (8) (that is, every system of curves with property I) into one of the same type are collineations and correlations.

That no other transformations are possible is seen as follows. If a contact transformation is to convert type (8) into itself, it must convert the part common to all systems of that type into itself. The curves defined by $y'' = 0$, that is, straight lines, obviously satisfy (8) for every form of G and H. It is obvious that no other (proper) curves satisfy all such equations. But since we are dealing with contact transformations and not merely point transformations, we must replace the concept *curve* by

the concept *union*. In the plane the only unions which are not (proper) curves are points. A point is regarded as made up of ∞^1 lineal elements; so x is constant, y is constant, y' is arbitrary, and therefore y'' and y''' are infinite. Point unions are to be regarded then as solutions of all equations (8). The common part thus consists of the ∞^2 straight lines and the ∞^2 points of the plane. If this is to go into itself, either points go into points and lines into lines, or else points go into lines and lines into points. We thus obtain only collineations and correlations.

That the collineations actually leave type (8) unchanged is easily verified analytically.* The work for correlations is simplified by observing that every correlation may be reduced, by means of collineations, to the form of Legendre's transformation

$$(9) \qquad x_1 = -y', \qquad y_1 = xy' - y, \qquad y_1' = -x,$$

(which is simply polarity with respect to the conic $x^2 + 2y - 1 = 0$). Extending (9), we find

$$(9') \qquad y_1'' = \frac{1}{y''}, \qquad y_1''' = \frac{y'''}{y''^3}.$$

This converts equation (8) into one of the same form

$$(10) \qquad y_1''' = G_1(x_1, y_1, y_1')y_1'' + H_1(x_1, y_1, y_1')y_1''^3,$$

the new coefficient functions being related to the old as follows:

$$(10') \qquad \begin{aligned} G_1 &= H(-y_1', x_1y_1' - y_1, -x_1), \\ H_1 &= G(-y_1', x_1y_1' - y_1, -x_1). \end{aligned}$$

This completes the proof of the theorem stated on the previous page.

78. If we impose property II on the system (8), that is, if we consider the subclass in which

$$(11) \qquad H = \frac{3}{y' - \omega(x, y)},$$

* *Trans. Amer. Math. Soc.*, vol. 7 (1906), p. 420.

the correlations are no longer available. That collineations actually convert this subclass into itself is readily verified. The same is true for the still narrower class, characterized by properties I, II, and III, in which the differential equation is of the form (cf. page 13)

$$(12) \qquad (y' - \omega)y''' = \{\lambda y'^2 + \mu y' + \nu\}y'' + 3y''^2.$$

79. We pass now to the case of dynamical trajectories, defined by type (7), and state the fundamental result:

Collineations are the only contact transformations of the plane which convert every system of ∞^3 dynamical trajectories (belonging to an arbitrary positional field of force) into such a system.

The only possibilities here also are collineations and correlations. The former actually have the required property. The latter have not, as is seen by observing that the application of the Legendre transformation (9) to a dynamical equation (8) will result in a new equation, which, while still of the general form (8), will not usually be of the dynamical form.*

80. Systems of trajectories are characterized by the set of five geometric properties of page 10. Therefore projective transformation will convert any system of curves having these properties into a system having the same properties. So, in spite of the fact that the properties as stated involve metric ideas (osculating parabolas, angles, circles of curvature, etc.), the set is actually projectively invariant. It ought to be possible therefore to *restate the geometric characterization in projective language.*

We shall not attempt to carry out this idea completely, and merely restate properties I and II as follows:

Consider the ∞^1 trajectories passing through a given point O in a given direction whose slope is y'. For each of these trajectories construct the conic which has four-point contact at O and touches the line determined by two arbitrarily selected

* We see from (10′) that the coefficients G and H, which are rational with respect to y', are converted into coefficients which are not usually rational.

points* A and B (which remain fixed in the following statements); through A and B draw tangents to the conic (in addition to the fixed line) and join the points of contact. *The lines thus constructed, one for each of the* ∞^1 *trajectories, will form a pencil* (property I).

As the initial direction (that is y') varies about O, the vertex of the pencil just described will move along a straight line† *passing through O* (property II).

The other properties, especially the fifth, are much more complicated.

81. In conclusion we point out another way in which the projective group enters in dynamics. If an arbitrary point transformation

$$x_1 = \Phi(x, y), \qquad y_1 = \Psi(x, y)$$

is applied to the differential equations

$$\ddot{x} = \varphi(x, y), \qquad \ddot{y} = \psi(x, y),$$

defining motion under a purely positional force, the new differential equations, of the more general form

$$\Phi_x\ddot{x} + \Phi_y\ddot{y} + \Phi_{xx}\dot{x}^2 + 2\Phi_{xy}\dot{x}\dot{y} + \Phi_{yy}\dot{y}^2 = \varphi(\Phi, \Psi),$$

$$\Psi_x\ddot{x} + \Psi_y\ddot{y} + \Psi_{yy}\dot{y}^2 + 2\Psi_{xy}\dot{x}\dot{y} + \Psi_{yy}\dot{y}^2 = \psi(\Phi, \Psi),$$

will usually define a motion due to a positional force together with a force depending on the velocity $\dot{x}$, $\dot{y}$. If this latter force is to be absent the transformation will be affine, as already remarked (§ 74). If, instead, we demand that the latter force shall act in the direction of the velocity (and thus be in the nature of a resistance), we find that the transformation may be any collineation.

More generally, *projective transformations are the only point*

* In the original metric statements these are of course the circular points at infinity.

† The force direction will be determined projectively as the harmonic of this line with respect to the lines joining O to A and B.

transformations which leave invariant the type

$$\ddot{x} = \varphi(x, y) + \dot{x}R(x, y, \dot{x}, \dot{y}),$$
$$\ddot{y} = \psi(x, y) + \dot{y}R(x, y, \dot{x}, \dot{y}),$$

defining motion of a particle under any positional force together with any resistance term acting in the direction of motion.

§§ 82–91. Conformal Transformations

82. The importance of conformal transformation is well known in connection with the theory of the potential. Geometric inversion or transformation by reciprocal radii, for example, yields the method of electric images due to Sir William Thomson. In connection with dynamics, the importance of general conformal transformations has been emphasized by Larmor, Goursat, and Darboux.*

83. Consider any conformal representation of the points of two surfaces S and S_1. The first fundamental forms of the surfaces may be taken to be

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2,$$
$$ds_1^2 = \lambda(Edu^2 + 2Fdudv + Gdv^2),$$

where corresponding points have the same parameters u, v. *The principal theorem is that every natural system on one surface becomes by the conformal representation a natural system on the other.* This is obvious if we remember that natural systems are obtained by minimizing an integral in which the integrand is the element of length multiplied by any point function. Hence

The only point transformations (in any space) which convert every natural family into a natural family are the conformal.

84. Consider now the ∞^3 dynamical trajectories on S produced by a conservative field of force, the work function being W. These consist of ∞^1 natural families, one for each value of the

* Cf. the discussion in Routh, Dynamics of a Particle, Nos. 628–635 (method of inversion and conjugate functions).

constant of total energy h. It will be convenient to refer to the particular natural system produced in the given field W for a particular value h, as the family due to $W + h$.

The corresponding family on S_1 is due to

$$\frac{W + h}{\lambda}.$$

Hence the ∞^1 related natural families on S, found by varying h, go over by the conformal representation into ∞^1 natural families which are *not* usually related, that is, do not form the complete system of trajectories belonging to a conservative field. The only case in which the new families are related arises when

$$W = \lambda,$$

for then the new systems are due to the work function

$$W_1 = 1/\lambda.$$

We then reach the conclusion that *in any conformal representation (excluding the trivial homothetic case*) there is a unique conservative force whose complete system of ∞^3 dynamical trajectories is converted into the complete system of some (usually distinct) conservative force. The work function of the force in question is defined by the squared ratio of magnification,*

$$W = \lambda = \frac{ds_1^2}{ds^2}.$$

85. Similar statements may be made for brachistochrones. Every system of ∞^2 brachistochrones due to any work function and a given value of h of course becomes such a system, for any natural family may be regarded as a family of brachistochrones. But *there is only one complete system of ∞^3 brachistochrones which is converted into a complete system, namely, that defined by the work*

* It is obvious that in this case *every* complete system of trajectories becomes a complete system. The same holds for brachistochrones and catenaries.

function

$$W = 1/\lambda.$$

For any other work function the ∞^1 families of brachistochrones, due to $W + h$, become ∞^1 non-related natural families on S_1 due to

$$\lambda(W + h).$$

86. In the case of catenaries due to $W + h$, the ∞^1 usually non-related natural families corresponding on S_1 are due to

$$\frac{W + h}{\sqrt{\lambda}}.$$

Hence *the only complete system of catenaries which is turned into a complete system is defined by the work function**

$$W = \sqrt{\lambda}.$$

87. Consider, for example, the conformal representation of the plane

$$z = x + iy = re^{i\theta}$$

on the plane

$$z_1 = x_1 + iy_1 = r_1 e^{i\theta_1}$$

defined by

$$z_1 = z^n,$$

where n is neither 0 nor 1.

Here the squared ratio of magnification is

$$\lambda = r^{2(n-1)} = r_1^{\frac{2(n-1)}{n}}.$$

* The three physical cases mentioned may be included in one general discussion by considering the extremals of

$$\int v^m ds = \int (W + h)^{m/2} ds = \text{minimum};$$

when $m = 1$, we have least action and trajectories; when $m = -1$, least time and brachistochrones. For every value of m we obtain, by varying h, a system of ∞^3 curves. Cf. the general discussion of the systems S_k defined (for arbitrary fields) in Chapter IV.

Applying the theorems stated above, we find that the trajectories generated by

$$W = r^{2(n-1)}$$

go over into the trajectories of a new field

$$W_1 = r_1^{2\left(\frac{1}{n}-1\right)}.$$

For brachistochrones the corresponding fields are

$$W = r^{-2(n-1)}, \qquad W_1 = r_1^{-2\left(\frac{1}{n}-1\right)};$$

and for catenaries

$$W = r^{n-1}, \qquad W_1 = r_1^{\frac{1}{n}-1}.$$

The particular transformation $z_1 = z^2$, that is, $n = 2$, gives rise to simple fields. Stating the results in terms of the law of the central forces obtained, instead of the corresponding work functions, we have:

The trajectories of a central force varying as r (that is, the conics described about the center of force as center) become the trajectories of a central force varying as r_1^{-2} (that is, the conics described about the center of force as focus).

The brachistochrones of a central force varying as r^{-3} become the brachistochrones of a central force of constant intensity.

The catenaries of a central force of constant intensity become the catenaries of a central force varying as $r_1^{-3/2}$.

88. Returning to the general conformal representation, we observe that ∞^1 natural families forming a complete system of trajectories can never become a complete system of brachistochrones. For the trajectories on S due to $W + h$ become ∞^1 natural families on S_1, which, when regarded as brachistochrones, are due to $\lambda/(W + h)$; and there is no work function which reduces this expression to the form of a function of u, v plus a constant depending only on h. Thus for a given (non-homothetic) conformal transformation there is *one* system of trajectories

which is converted into a system of trajectories, and *one* system of brachistochrones which is converted into a system of brachistochrones, but there is *no* system of trajectories which is converted into a system of brachistochrones. The same is true for any two of the three types trajectories, brachistochrones, catenaries or of the infinite number of types described in the preceding footnote (page 83).

89. As another application, consider the *velocity curves* connected with a plane field of force whose work function is $W(x, y)$. For a given speed v_0, we obtain ∞^2 such curves, defined by the property that the curvature at each point and direction equals the curvature of a free particle starting out from that point and direction with the speed v_0. The differential equation of this velocity system is

$$y'' = \frac{(W_y - y'W_x)(1 + y'^2)}{v_0^2}.$$

This is recognized as a natural family; it corresponds to the geodesics of the surface whose first fundamental form is

$$e^{\frac{2W}{v_0^2}}(dx^2 + dy^2).$$

By varying v_0 we obtain the ∞^1 velocity systems belonging to the given field; they are pictured by the geodesics of ∞^1 surfaces.

Consider now a conformal representation of the xy-plane upon itself. This converts $dx^2 + dy^2$ into

$$e^H(dx^2 + dy^2),$$

where $H(x, y)$, by known theory, is a harmonic function. We thus obtain ∞^1 new natural families corresponding to the geodesics of the ∞^1 surfaces

$$e^{\frac{2}{v_0^2}W + H}(dx^2 + dy^2).$$

These ∞^1 natural families cannot usually be regarded as related velocity systems for some new field: the requisite condition is that W shall be the same as H except for a constant factor.

Hence for a given conformal transformation of the plane (which is not merely a similitude), there is a unique complete velocity system belonging to a conservative field of force which is converted into a complete system. The unique work function is

$$W = H = \log \lambda,$$

where λ denotes the squared ratio of magnifaction in the given conformal representation. The fields obtained are *Laplacian*, that is, satisfy the condition

$$W_{xx} + W_{yy} = 0.$$

As an example, the transformation $z_1 = \log z$ converts the ∞^3 velocity curves of the field $W = \log r$ (in which the force varies inversely as the distance from the origin) into the ∞^3 velocity curves of the field $W_1 = x_1$ (force vertical and constant).

90. It was shown above that conformal transformations are the only point transformations which convert every natural family into a natural family. Natural families are characterized by properties A and B of § 31. It is of interest to notice that property A by itself is conformally invariant. The most general system having this property (that osculating circles constructed at any point have another point in common) is what we have termed a velocity system. We now prove that

The only point transformations which convert every velocity system into a velocity system are the conformal transformations.

Consider say the three-dimensional case, where the general velocity system is

$$y'' = (\psi - y'\varphi)(1 + y'^2 + z'^2), \qquad z'' = (\chi - z'\varphi)(1 + y'^2 + z'^2).$$

The only curves which are common to all such systems must

satisfy

$$1 + y'^2 + z'^2 = 0, \qquad y'' = 0, \qquad z'' = 0,$$

and are therefore the minimal straight lines of space. Since the only transformations converting minimal lines into minimal lines are conformal, we have the result stated. That conformal transformations actually leave the velocity type invariant is easily verified analytically*. The result is obvious synthetically (in the case of more than two dimensions) since the conformal group converts circles into circles and bundles of circles into bundles. Hence if the original system possesses property A, the same will be true of the transformed system.

91. It may be shown that, for any given non-conformal transformation, there exists one and only one velocity system which is converted into a velocity system.

§§ 92–94. Contact Transformations

92. With each natural family, or, what is the same, with each isotropic medium, there is associated a definite infinitesimal contact transformation. This connection, which appears implicitly in Hamilton's fundamental memoir of 1835, was worked out in detail by S. Lie.†

If the index of refraction is $\nu(x, y, z)$, the associated contact transformation has the characteristic function

$$(1) \qquad \nu(x, y, z)\sqrt{1 + p^2 + q^2},$$

where x, y, z, p, q are considered as the coordinates of a surface element. If the one-parameter group generated is applied to an arbitrary surface the resulting ∞^1 surfaces form a wave set. The trajectories or rays appear as the path curves of this group. Lie showed that the category of transformations which thus

* Cf. *American Journal of Mathematics*, vol. 27 (1906), p. 213, for the two-dimensional case.

† "Die infinitesimalen Berührungstransformationen der Mechanik," *Leipziger Berichte* (1889), pp. 145–153. A very elegant discussion, with new results, is given by Vessiot, *Bull. Soc. math. de France*, vol. 34 (1906), pp. 230–269.

appears, with a characteristic function of type (1), and which he termed "the infinitesimal contact transformations of mechanics," is distinguished geometrically by the fact that the so-called* *transversality* relation reduces to orthogonality.

93. The following simple and easily proved theorem appears to be new.

The alternant (or Klammerausdruck of Lie) of the contact transformations associated with any two media is always a point transformation.

94. Here we are dealing with two natural families in the same three-dimensional space. In connection with the most general problem of dynamics (page 70), spaces of any dimensionality must be considered, with arbitrary variable curvature. The space depends on the quadratic form defining the kinetic energy: this determines the quadratic expression appearing under the radical in the generalization of (1). The potential† determines the factor ν which may be any point function. The general theorem is then as follows:

The alternant of the contact transformations associated with two dynamical problems (or natural families) will be a point transformation when, and only when, the two expressions for the kinetic energy are either the same or differ by a factor (which may be any point function); the two potential energies remain entirely arbitrary.

In particular, if any two natural families are constructed in the same space (which space is entirely arbitrary), the alternant will be a point transformation.

For a detailed discussion of the two-dimensional case, including a number of converse results, the reader is referred to the author's paper, cited in the first footnote below.

* Lie does not use this term. The author borrows it from the closely connected problem in the calculus of variation. See "The infinitesimal contact transformations of mechanics," *Bull. Amer. Math. Soc.*, vol. 16 (1910), pp. 408–412.

† Here considered as including the energy constant h, which is fixed, since we are dealing with a natural family.

§§ 95–97. A Group of Space-time Transformations

95. In the fundamental transformation of the relativity theory, known as the Lorentz transformation, the position coordinates x, y, z and the time coordinate t are merged: the new position and the new time appear as functions of both the original position and the original time. The Lorentz group is composed of the linear transformations of the four variables x, y, z, t which leave invariant the quadric

$$x^2 + y^2 + z^2 - c^2t^2 = 0.$$

Its importance is due to the fact that it leaves unaltered the form of the Maxwell equations.

We consider in this section an entirely different group of space-time transformations, depending on arbitrary functions instead of arbitrary constants. It arises in connection with ordinary (newtonian) dynamics in the theory of forces depending on the time as well as position.

We confine the discussion for the sake of simplicity to the case of two dimensions. What transformations of the three variables x, y, t will convert any set of equations of the form

$$(1) \qquad \frac{d^2x}{dt^2} = \varphi(x, y, t), \qquad \frac{d^2y}{dt^2} = \psi(x, y, t)$$

into another set of the same form? An arbitrary transformation would produce equations representing a force depending, not only on x, y, t, but also on the velocity dx/dt, dy/dt. The problem is to find those peculiar transformations which do not introduce the velocity in the final equations. The result is as follows:

The only space-time transformations which convert every space-time field of force into a space-time field are those of the form

$$(2) \qquad t_1 = f(t), \qquad x_1 = (ax + by)\sqrt{f'(t)} + g(t),$$
$$y_1 = (cx + dy)\sqrt{f'(t)} + h(t).$$

The group thus involves three arbitrary functions $f(t)$, $g(t)$, $h(t)$ as well as four arbitrary constants a, b, c, d.

96. Another representation of the same group, which has the advantage of avoiding radicals, is

$$
(3)\qquad \frac{dt_1}{dt} = [\lambda(t)]^2, \qquad \begin{aligned} x_1 &= (ax + by)\lambda(t) + \mu(t), \\ y_1 &= (cx + dy)\lambda(t) + \nu(t). \end{aligned}
$$

When such a transformation is applied to equations (1), the new equations are found to be

$$
\lambda^5 \ddot{x}_1 = (\lambda\ddot{\lambda} - 2\dot{\lambda}^2)(ax + by) + \lambda^2(a\varphi + b\psi) + \lambda\ddot{\mu} - 2\dot{\lambda}\mu,
$$
$$
\lambda^5 \ddot{y}_1 = (\lambda\ddot{\lambda} - 2\dot{\lambda}^2)(cx + dy) + \lambda^2(c\varphi + d\psi) + \lambda\ddot{\nu} - 2\dot{\lambda}\nu.
$$

Of course the original variables x, y, t are here to be replaced by their values in the new variables x_1, y_1, t_1.

97. The transformation converts the space-time curves of the original force into the space-time curves of a new force. Of course it is not a point transformation of the xy-plane, so it does not, as was the case for the Appell transformation (page 76), convert trajectories into trajectories. These remarks apply even in the special case where the force is positional. Consider, as a simple example, the transformation

$$
t_1 = \tfrac{1}{2}e^{2t}, \qquad x_1 = xe^t, \quad y_1 = ye^t,
$$

applied to the equations

$$
\ddot{x} = x, \qquad \ddot{y} = y.
$$

The transformed equations are found to be

$$
\ddot{x}_1 = 0, \qquad \ddot{y}_1 = 0.
$$

The first field is central, the force varying directly as the distance, so that the trajectories are ∞^3 conics with the same center. The second force is everywhere zero, so the trajectories are merely ∞^2 straight lines.

CHAPTER IV

CONSTRAINED MOTIONS IN A FIELD. GENERALIZATION OF THE TRAJECTORY PROBLEM INCLUDING BRACHISTOCHRONES AND CATENARIES

§§ 98–114. SYSTEMS S_k DEFINED BY $P = kN$

98. In connection with a field of force, the only curves usually studied are the lines of force and the trajectories. In the plane the lines of force form a simply infinite system, and the trajectories a triply infinite system. The former system has no peculiar properties, since any set of ∞^1 curves may be regarded as the lines of force in some field, in fact in an infinite number of different fields. The triply infinite system of trajectories has peculiar properties which have been discussed in Chapter I. Other noteworthy systems of curves are connected with the field, for example, brachistrochrones, catenaries, velocity curves, and tautochrones.

99. Omitting the tautochrones, *the other three systems named, together with the trajectories, may all be obtained as special cases of this simple general problem*: to find curves along which a constrained motion is possible such that the pressure is proportional to the normal component of the force.

100. If an arbitrary curve is drawn in the plane field of force, and the particle, of say unit mass, is started along it from one of its points with a given speed, the constrained motion along the given curve is determined. The acceleration along the curve is given by T, the tangential component of the force vector. So the speed at any point is determined by

$$v^2 = \int T ds. \tag{1}$$

The pressure P (of course normal to the curve, since the curve

is considered smooth) is given by the elementary formula

$$(2) \qquad P = \frac{v^2}{r} - N.$$

If we increase the initial speed, the effect is to increase v^2 by a constant c; and hence P changes by the addition of a term of the form c/r.

101. If the given curve is a trajectory, the initial speed may be so chosen that the pressure vanishes throughout the motion; that is, trajectories may be defined as curves of no constraint. Of course, if a different initial speed is used, P will be of the form c/r; but, as regards the curves, they are completely characterized by $P = 0$.

102. If the given curve is a brachistochrone and if the motion along it is brachistochronous, Euler proved (assuming the force to be conservative) that the pressure was double the normal component of the acting force and opposite to it in direction, that is, $P = -2N$. If the force is not conservative, the real brachistochrones, as defined by a problem of the calculus of variations, form a quadruply infinite system. The curves defined by the property $P = -2N$ then form a triply infinite system of what should be called pseudo-brachistochrones. These curves are really brachistochrones only in the conservative case. No ambiguity however will arise by terming the system here considered brachistochrones instead of pseudo-brachistochrones.

103. *The general problem suggested is to find curves such that P shall be proportional to N.* So $P = kN$. To a given value of k there correspond ∞^3 such curves; the system so obtained will be denoted by S_k. *The four special cases of physical interest are as follows:*

$k = 0$ gives S_0, the system of *trajectories*;

$k = -2$ gives S_{-2}, the system of *brachistochrones*;

$k = 1$ gives S_1, the system of *catenaries*;

$k = \infty$ gives S_∞, the system of *velocity curves*.

104. The last case requires a justification in terms of limits which is easily carried out analytically.

105. The third case follows from the known fact that when an inextensible flexible homogeneous string is suspended in any field of force, the resulting form of equilibrium, called a catenary in the general sense of the term, has the dynamical property that when a particle, started out with the proper initial velocity, rolls along the curve, the pressure at any point equals the normal component of the force: that is, catenaries are defined by $P = N$, corresponding to $k = 1$.

106. Of course a triply infinite system S_k exists for any value of the parameter k. The differential equation of the system, in intrinsic form, is easily obtained by eliminating v from the equations

(3) $$v^2/r = (k+1)N, \qquad vv_s = T.$$

The result is

(4) $$Nr_s = (n+1)T - r\mathfrak{N},$$

where

(4′) $$n = 2/(k+1).$$

We may readily find various properties from this intrinsic equation, but in order to obtain a complete set it is necessary to have recourse to the equivalent equation in cartesian coordinates

(5) $$(\psi - y'\varphi)y''' = \{\psi_x + (\psi_y - \varphi_x)y' - \varphi_y y'^2\}y'' - \left\{3 + \frac{(n-2)(\varphi + y'\psi)}{1+y'^2}\right\}y''^2.$$

This obviously reduces to the familiar trajectory equation of §1 when $n = 2$, corresponding to $k = 0$. Brachistochrones correspond to $n = -2$, catenaries to $n = 1$, velocity curves to $n = 0$.

107. We now state the characteristic properties of a system of the above type for any value of n, that is, any value of k.

is considered smooth) is given by the elementary formula

$$(2) \qquad P = \frac{v^2}{r} - N.$$

If we increase the initial speed, the effect is to increase v^2 by a constant c; and hence P changes by the addition of a term of the form c/r.

101. If the given curve is a trajectory, the initial speed may be so chosen that the pressure vanishes throughout the motion; that is, trajectories may be defined as curves of no constraint. Of course, if a different initial speed is used, P will be of the form c/r; but, as regards the curves, they are completely characterized by $P = 0$.

102. If the given curve is a brachistochrone and if the motion along it is brachistochronous, Euler proved (assuming the force to be conservative) that the pressure was double the normal component of the acting force and opposite to it in direction, that is, $P = -2N$. If the force is not conservative, the real brachistochrones, as defined by a problem of the calculus of variations, form a quadruply infinite system. The curves defined by the property $P = -2N$ then form a triply infinite system of what should be called pseudo-brachistochrones. These curves are really brachistochrones only in the conservative case. No ambiguity however will arise by terming the system here considered brachistochrones instead of pseudo-brachistochrones.

103. *The general problem suggested is to find curves such that P shall be proportional to N.* So $P = kN$. To a given value of k there correspond ∞^3 such curves; the system so obtained will be denoted by S_k. *The four special cases of physical interest are as follows:*

$k = 0$ gives S_0, the system of *trajectories*;

$k = -2$ gives S_{-2}, the system of *brachistochrones*;

$k = 1$ gives S_1, the system of *catenaries*;

$k = \infty$ gives S_∞, the system of *velocity curves.*

104. The last case requires a justification in terms of limits which is easily carried out analytically.

105. The third case follows from the known fact that when an inextensible flexible homogeneous string is suspended in any field of force, the resulting form of equilibrium, called a catenary in the general sense of the term, has the dynamical property that when a particle, started out with the proper initial velocity, rolls along the curve, the pressure at any point equals the normal component of the force: that is, catenaries are defined by $P = N$, corresponding to $k = 1$.

106. Of course a triply infinite system S_k exists for any value of the parameter k. The differential equation of the system, in intrinsic form, is easily obtained by eliminating v from the equations

$$(3) \qquad v^2/r = (k+1)N, \qquad vv_s = T.$$

The result is

$$(4) \qquad Nr_s = (n+1)T - r\mathfrak{N},$$

where

$$(4') \qquad n = 2/(k+1).$$

We may readily find various properties from this intrinsic equation, but in order to obtain a complete set it is necessary to have recourse to the equivalent equation in cartesian coordinates

$$(5) \qquad \begin{aligned}(\psi - y'\varphi)y''' = \{\psi_x + (\psi_y - \varphi_x)y' - \varphi_y y'^2\}y'' \\ - \left\{3 + \frac{(n-2)(\varphi + y'\psi)}{1+y'^2}\right\}y''^2.\end{aligned}$$

This obviously reduces to the familiar trajectory equation of §1 when $n = 2$, corresponding to $k = 0$. Brachistochrones correspond to $n = -2$, catenaries to $n = 1$, velocity curves to $n = 0$.

107. We now state the characteristic properties of a system of the above type for any value of n, that is, any value of k.

Characteristic Properties of the System S_k

Property 1.—For any given element (x, y, y') the foci of the osculating parabolas of the single infinity of curves determined by the given element lie on a circle passing through the given point.

Property 2.—At any point O the tangent of the angle which the focal circle makes with the given element is to the tangent of the angle which the given element makes with a certain direction fixed at O (the direction of the acting force) as 3 is to $n + 1$, that is, as $3k + 3$ is to $k + 3$.

Property 3. Through a given point there pass a single infinity of curves admitting hyperosculating circles of curvature; the centers of these circles lie on a conic passing through the given point in the direction of the force vector.

Property 4.—The normal at the given point O cuts the conic described in property 3, at a distance equal to $n + 1$, that is $(k + 3)/(k + 1)$, times the radius of curvature of the line of force passing through O.

Property 5.—This is of the same form as property V (§ 3) obtained in the discussion of trajectories, the number 3 being replaced by the number $n + 1$. In the notation of page 11

$$\frac{\partial}{\partial x}\frac{1}{AA'} + \frac{\partial}{\partial y}\frac{1}{BB'} + \frac{\omega\omega_{xy} - \omega_x\omega_y}{(n+1)\omega^3} = 0.$$

108. The special case where n equals -1, that is, the system S_{-3}, is exceptional and requires a separate discussion; but as we do not need the results, this case is omitted.

109. While the properties corresponding to different values of k are analogous, they are of course not identical. The first property is common to all the systems. But the second property involves the parameter k. Thus, while for trajectories the constant ratio that appears is 1 (bisection), it is -3 for brachistochrones, 3/2 for catenaries, and 3 for velocity curves. Not only are the triply infinite systems S_k, corresponding to different values of k, distinct in any given field of force, but also no two

systems arising in two distinct fields can ever coincide. For example, if a certain system of ∞^3 curves arises as trajectories in one field, it cannot also arise as catenaries in either the same or another field.

110. If we combine all the systems S_k, in a given field (φ, ψ), we obtain a quadruply infinite system which we now proceed to study. The differential equation of the fourth order defining this system is readily obtained by eliminating k from the equation of S_k. It is more convenient to carry this out in terms of intrinsic quantities, using either the radius of curvature and its first and second derivatives with respect to the arc, quantities denoted by r, r_s, r_{ss}, or else the radius of curvature together with the radii of the first and second evolute, quantities which we denote by r, r_1, r_2. The two sets of quantities are equivalent, being connected by the relations $r_1 = rr_s$, $r_2 = r^2 r_{ss} + rr_s^2$. The equation of the quadruply infinite system may then be put, using the notation of § 2, into the form

$$\begin{vmatrix} Nr_s + r\mathfrak{N} & T \\ Nr_{ss} + \left(2\mathfrak{N} - \dfrac{T}{r}\right) r_s & \mathfrak{T} + \dfrac{N}{r} \end{vmatrix} 0 = .$$

This may be written in either of the forms

$$r_{ss} = (\beta_1 + \beta_2 r^{-1}) r_s + \beta_3 r + \beta_4,$$
$$r_2 = r^{-1} r_1^2 + (\beta_1 r + \beta_2) r_1 + \beta_3 r^3 + \beta_4 r^2,$$

where the β's are functions of x, y, y'.

111. We notice first that r_2 is quadratic with respect to r_1. Hence for given values of x, y, y', r, that is for a given curvature element, the ∞^1 curves of the system have the property that the locus of the third center of curvature is a parabola with axis parallel to the fixed radius of curvature, that is, perpendicular to the initial direction y'.

112. An equivalent statement is this: If for each of the curves we construct the osculating conic (five-point contact), the locus

of the centers of these conics is a conic passing through a given point in the given direction. It is perhaps worth while to restate this, so far as it concerns the four special cases of physical interest, as follows: In any plane field of force select any fixed element of curvature; corresponding to the initial values of x, y, y' and r so given, construct the unique trajectory, unique brachistochrone, unique catenary, the unique velocity curve, and the respective centers of the osculating conics; the four centers so found and the given point (x, y) will lie on a conic passing through the latter point in the given direction y'. (Cf. the first footnote on page 98.)

113. Keeping the curvature element fixed and varying the parameter k, the value of r, or, what is equivalent, of r_1, varies linearly. As above, let n denote the fraction $2/(k+1)$; then if values of n forming an arithmetic progression are selected, the corresponding values of r_1 also form an arithmetic progression. The successive differences in the values of r_1 corresponding to the case of trajectory, brachistochrone, catenary, and velocity curve are proportional to 4, -3, 1.

114. If in the system S_k we keep x, y, y' fixed and vary r, two limiting cases of interest arise. First, if r becomes infinite, then r_n is also infinite, and the limiting curve obtained is a straight line. In fact the ∞^2 straight lines of the plane form part of every system S_k.

On the other hand, if r approaches zero, then r_n approaches a definite limit

$$(n+1)T/N.$$

Remembering that the tangent of the angle of deviation is one third of r_n, we may state the result obtained as follows: In any system S_k if we take any lineal element and let r approach zero, the tangent of the corresponding angle of deviation is to the tangent of the angle which the force vector makes with the normal to the given element in the fixed ratio of $n+1$ to 3. The special values of this ratio for the four special systems of physical interest are respectively 1, $-1/3$, $2/3$, $1/3$. In the case of trajectories, it is noteworthy that the limiting position of the axis

of deviation coincides with the direction of the force acting at the given point.

§§ 115–116. Curves of Constant Pressure

115. We now consider a second simple generalization of the problem $P = 0$, defining trajectories. We consider, namely, curves corresponding to $P = c$, where c denotes any constant. The curves obtained may be termed curves of constant pressure: only along such a curve is a constrained motion of a particle possible such that the pressure against the curve remains constant.

For a given value of c a system of ∞^3 such curves is obtained, whose intrinsic equation, found by differentiating the relation

$$P \equiv v^2/r - N = c,$$

is

$$(c + N)r_s = 3T - \mathfrak{N}.$$

We see that this system for any value of c retains property I of the system of trajectories. Omitting the discussion of the higher properties of these triply infinite systems we consider the quadruply infinite system whose differential equation, found by eliminating c, may be written in either of the intrinsic forms

$$(\mathfrak{N}r^2 - 3Tr)r_{ss} = (2r\mathfrak{N} - T)r_s^2 + [\mathfrak{N}_1 r^2 + (\mathfrak{N}_2 - 3\mathfrak{T})r - 3N]r_s,$$

$$r(\mathfrak{N}r - 3T)r_2 = (3r\mathfrak{N} - 4T)r_1^2 + [\mathfrak{N}_1 r^2 + (\mathfrak{N}_2 - 3\mathfrak{T})r - 3N]rr_1.$$

This gives the totality of ∞^4 curves of constant pressure defined by a given field.

As regards special cases of interest, we note, in addition to $c = 0$, giving trajectories, the case $c = \infty$ which gives $r_s = 0$, defining circles; hence for any field of force the ∞^4 curves of constant pressure include the ∞^3 circles of the plane, which arise in fact as curves of infinite pressure.

116. The quadruply infinite system which here arises, as well as that obtained in the previous problem $P = kN$, comes under

the category represented by a differential equation of the type*

$$y^{\mathrm{IV}} = Ay'''^2 + By''' + C.$$

It therefore enjoys the property, previously stated in the other problem (§ 112), that the locus of the centers of the osculating conics corresponding to any element (x, y, y', y'') is a conic touching the element (x, y, y'). Of course, since the forms of A, B, C in the two problems are quite distinct, the systems are distinguished in their higher properties.

§§ 117 118. Tautochrones

117. Tautochrones are not included in either of the previous problems. They are not distinguished by any simple law of pressure.† The condition for a tautochrone is that the resulting constrained motion of a particle along the curve be harmonic, that is,

$$T = k(s - s_0), \tag{1}$$

where k is a constant (which is negative for actual and positive for virtual tautochrones) and $s - s_0$ denotes the arc reckoned from a fixed point of the curve, the center of the tautochronous motion. From this

$$T_{ss} = 0 \tag{2}$$

and hence, by expansion, *the general equation of the system of ∞^3 tautochrones in any field is*‡

$$Nr_s = \mathfrak{T}_1 r^2 + (\mathfrak{T}_2 + \mathfrak{N})r - T, \tag{3}$$

where the notation is that of § 2.

* This type (noteworthy in that it unifies many distinct mathematical and physical problems) first presented itself in the author's study of "Systems of extremals in the calculus of variations," *Bull. Amer. Math. Soc.*, vol. 13 (1907), p. 290: the extremals of any integral of the second order $\int f(x, y, y', y'')dx$ form a system of that type. In these lectures other physical problems leading to species included in this type are treated in §§ 110, 135, 137.

† It may be shown that during any tautochronous motion

$$P = k(s - s_0)^2/r - N.$$

‡ "Tautochrones and brachistochrones," *Bull. Amer. Math. Soc.*, vol. 15 (1909), pp. 475–483.

We see that r_s is a quadratic function of r, and not a linear function as in the case of trajectories and the other systems S_k. For a discussion of the geometric properties of tautochrones, we refer to the dissertation of H. W. Reddick.*

118. There is no field in which the tautochrones coincide with the trajectories, or with any of the systems S_k, in either the same or some other field, except for the case $k = -2$ corresponding to brachistochrones. The classical work of Huygens and J. Bernoulli showed that for a uniform field the system of tautochrones is identical with the system of brachistochrones. The author has shown that the only other field where such duplication occurs is that in which the force is central and varies directly as the distance. The only case of duplication in two distinct fields is as follows: The tautochrones of the field $\varphi = 0,\ \psi = y$ coincide with the brachistochrones of the field $\varphi = 0, \quad \psi = y^{-3}$. The particular fields arising in this duplication problem are included in the interesting class of fields, involving eight parameters, characterized by the vanishing of the element function T_1. For such a field r_s, according to (3), becomes linear in r, and hence the ∞^2 straight lines of the plane are included in the system of tautochrones.†

118′. Each of the ∞^3 tautochrones in a given field has associated with it a certain time of oscillation, determined by the value of the constant k in (1). To each value of the period, that is, to each value of k, corresponds a certain family of ∞^2 tautochrones, whose differential equation, in implicit form, is

$$r(k - \mathfrak{T}) = N,$$

or, expanded,

$$(\psi - y'\varphi)y'' = k(1 + y'^2) - \{\varphi_x + (\varphi_y + \psi_x)y' + \psi_y y'^2\}.$$

We pass over the easy geometric interpretation; and note merely the special family, corresponding to the value $k = 0$, for which

* *Amer. Jour. of Math.*, vol. 33 (1911).

† The corresponding problem in space is treated in Reddick's paper and gives a class of fields involving twenty parameters.

the period is infinite. This separates the actual from the virtual tautochrones.

§ 119. Non-uniform Catenaries

119. It is a familiar fact that vertical parabolas appear in elementary dynamics in two distinct dimensions; first, as trajectories of a cannon ball, and secondly as forms of equilibrium of a chain in which the mass (or load) of any element is proportional to the horizontal projection of that element. Here the force is ordinary gravity. The question arises whether any other fields of force give rise to a like duplication.

We first consider the following general problem of non-uniform catenaries. If a flexible string or chain, in which the mass of any element of length is proportional to some given function μ of x, y, y', is suspended in a positional field, the possible forms of equilibrium are defined by the equation

$$Nr_s = 2T - (1 + y'^2)\{N\mu_y - r\}\Re + (1 + y'^2)^{\frac{1}{2}}N(\mu_x + y'\mu_y)\}.$$

This represents the ∞^3 non-uniform catenaries for a given field $\varphi(xy)$, $\psi(xy)$ and a given density law $\mu(x, y, y')$, where μ denotes $\log \mu$.

On the other hand, the trajectories in the given field are defined by the equation

$$Nr_s = 3T - r\Re.$$

Our problem then is to find those fields for which the two systems described coincide. *The result obtained is that the field must be central or parallel.* The detailed result is as follows:

In any central field of force the ∞^3 trajectories may be also obtained as catenaries by loading the chain so that its density is proportional to the perpendicular dropped from the center to the tangent line. In the more special case where the field is parallel, the density is proportional to the sine of the angle between the element of the curve and the force.

It is easy to obtain analogous comparisons between brachistochrones and catenaries. In this case the density must vary inversely as the cube of the perpendicular dropped from the center (or of the sine of the angle referred to above). For example, in the case of gravity the vertical cycloids which appear as brachistochrones may be obtained as catenaries by causing the load applied to any element to vary inversely as the cube of its horizontal projection.

All the results may be included in a generalization found by comparing the non-uniform catenaries with the systems denoted by S_k in § 103. The density must vary as the $(n-1)$th power of the perpendicular, where n is the number defined on page 93. The field is necessarily central or parallel.

CHAPTER V

MORE COMPLICATED TYPES OF FORCE

§§ 120–122. Motion in a Resisting Medium

120. We consider the motion of a particle moving in the plane under a positional field of force and influenced by a resisting medium, the resistance acting in the direction of the motion and varying as some function of the speed v. The equations of motion will then be of the form

$$(1) \qquad \ddot{x} = \varphi(x, y) + \dot{x} f(v), \qquad \ddot{y} = \psi(x, y) + \dot{y} f(v),$$

where the resistance R is equal to

$$R = v f(v).$$

The differential equation of the trajectories is found to be

$$(2) \qquad \begin{aligned}(\psi - y'\varphi)y''' = \{\psi_x + y'(\psi_y - \varphi_x) - y'^2\varphi_y\} \\ - 3\varphi y''^2 - 2f\sqrt{\psi - y'\varphi}\, y''^{3/2},\end{aligned}$$

where the argument v of f is to be expressed in terms of x, y, y', y'' by means of

$$v^2 = \frac{(\psi - y'\varphi)(1 + y'^2)}{y''}.$$

Consider now the ∞^1 trajectories starting from a given element (x, y, y'). The focal locus, that is, the locus of the foci of the osculating parabolas, varies in shape with the function f, that is, with the law of resistance.

We know that, if there is no resistance, property I of § 3 holds, that is, the focal locus is a circle passing through the given point. Are there any resisting media for which this property is preserved? A simple discussion shows that there are, the appro-

priate media being those for which R is of the form $Av^2 + B$.

For such media, property II will not usually be fulfilled; in fact *the only medium preserving the properties I and II is that in which the resistance varies as the square of the speed.*

If we impose also property III, both A and B must vanish, that is, the resistance vanishes and the force is purely positional.

It is of interest to examine the case where the resistance varies as any power v^n of the speed. The differential equation of the trajectories is then of the form

$$y''' = ay'' + by''^2 + cy''^m,$$

where

$$m = \tfrac{1}{2}(4 - n).$$

The focal locus is a curve whose inverse with respect to the given point is

$$X = a_1 + b_1 Y + c_1 Y^{m-1}.$$

This becomes a straight line (as in the case of no resistance), when m is 1 or 2, that is, when n is 2 or 0.

The curve is a conic when m is 3 or 0 or 3/2, that is, when n has one of the values -2 or 4 or 1. When $n = -2$ the conic is a parabola with its axis parallel to the given element. When $n = 4$ it is a hyperbola, asymptotic to the line of the given initial element. When $n = 1$ it is a parabola touching the initial line (not at the given point).

121. We now state briefly the corresponding results in ordinary space. No matter what the law of resistance is, property I (of the set of four properties for space given in § 11) is fulfilled; for the osculating planes necessarily pass through the force vector. The only laws for which property II is preserved are those included in

$$R = Av^2 + B.$$

If property III is also to be preserved, the resistance must vanish.

122. The results may be derived easily from the intrinsic equations

$$(3) \qquad v^2 = rN, \qquad vv_s = T + R,$$

obtained by taking components of the acting forces along the normal and tangent to the trajectory. The geometric equation, resulting from the elimination of v, is of the form*

$$(4) \qquad Nr_s = -r\mathfrak{R} + 3T + 2R.$$

This gives the relation between r_s (the rate of variation of r with respect to s) and r (the radius of curvature). The resistance R, which is given as a function of v, is here to be expressed in terms of r by means of the first of the relations (3). If property I, of plane trajectories, is to hold, r_s must be a linear integral function of r; this will be the case not only when R vanishes, but also, as stated above, when it is of the form $Av^2 + B$.

§§ 123–126. Particle on a Surface

123. The motion of a particle on any constraining surface

$$x = \varphi(u, v), \qquad y = \psi(u, v), \qquad z = \chi(u, v)$$

under any positional forces may be investigated most simply by means of the Lagrangian equations

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot u}\right) - \frac{\partial T}{\partial u} = U, \quad \frac{d}{dt}\left(\frac{\partial T}{\partial \dot v}\right) - \frac{\partial T}{\partial v} = V,$$

where T is the kinetic energy

$$2T = E\dot u^2 + 2F\dot u\dot v + G\dot v^2$$

and U, V are the components of the force given as functions of u, v.† The explicit equations of motion are of the form

$$\ddot u = \Phi + A_0\dot u^2 + 2A_1\dot u\dot v + A_2\dot v^2,$$
$$\ddot v = \Psi + B_0\dot u^2 + 2B_1\dot u\dot v + B_2\dot v^2;$$

* From this we may obtain the following dynamical result: If a particle starts from rest, the initial radius of curvature of the trajectory is to the radius of curvature of the line of force passing through the initial point as $3T + 2R$ is to T. When R vanishes we have the simple result previously stated.

† See for example Whittaker, Analytical Dynamics, p. 390, and Hadamard, *Jour. de Math.* (5), vol. 3, p. 331.

where Φ, Ψ define the force and the A's and B's are functions of u, v depending only on the given surface.

124. We observe that here $\ddot{u}$, $\ddot{v}$ depend not only on the position u, v but also upon the velocity $\dot{u}$, $\dot{v}$. Hence the motion in the uv-plane corresponding to the actual motion on the surface is not usually generated by any positional force in that plane. The only exception arises when the A's and the B's vanish identically: this is the case only if the given surface is developable, and if its representation on the uv-plane differs from its development on the plane by at most an affine transformation.

Another problem including this as a special case is to determine when the motion in the uv-plane can be regarded as due to a positional force together with a resistance acting in the direction of the motion. The condition for this is

$$\frac{A_0\dot{u}^2 + 2A_1\dot{u}\dot{v} + A_2\dot{v}^2}{B_0\dot{u}^2 + 2B_1\dot{u}\dot{v} + B_2\dot{v}^2} = \frac{\dot{u}}{\dot{v}}.$$

Expanding, we find four conditions on the six functions A, B, which turn out to be precisely the conditions that the geodesics of the surface shall be pictured by straight lines, a result which may be proved directly. Hence the only case in which the motion on the surface is pictured in the uv-plane by a motion due to a positional force together with a resistance depending on the velocity components and acting in the direction of the motion, is that in which the surface has constant curvature and the representation is geodesic.

125. We proceed with the general equations of motion. If we eliminate the time, we obtain the differential equation of the third order defining the ∞^3 trajectories in the form

$$(\Psi - v'\Phi)v''' = \{\delta_0 + \delta_1 v' + \delta_2 v'^2 + \delta_3 v'^3 + \delta_4 v'^4 + \delta_5 v'^5\} \\ + \{\epsilon_0 + \epsilon_1 v' + \epsilon_2 v'^2 + \epsilon_3 v'^3\}v'' - 3\Phi v''^2,$$

where the coefficients are functions of u, v. We confine ourselves to the observation that the picture curves in the uv-

plane come under the type

$$v''' = F_0 + F_1 v'' + F_2 v''^2,$$

where the coefficients are lineal-element functions: the focal locus is thus not a circle, but a special quartic. Hence if we consider the ∞^1 trajectories on the surface obtained by starting a particle at a given point in a given direction with different speeds, the picture curves in the uv-plane have osculating parabolas at the common point whose foci lie on a special quartic curve.

126. What is the simplest property of the actual trajectories described on the surface? What is, in particular, the locus of the osculating spheres of the ∞^1 trajectories considered?

To answer this we take our surface not in parametric form, but in the explicit form

$$z = f(x, y).$$

We may take the given point as origin, the tangent plane as the xy-plane, and the fixed initial direction as that of the axis of x. We find, by differentiating the equation of the surface and making use of $y' = 0$, $z' = 0$, that

$$z'' = a, \qquad z''' = b + cy'',$$

where a, b, c are constants, equal respectively to the values of the partial derivatives f_{xx}, f_{xxx}, $3f_{xy}$ at the origin. Again, from the general equation of the trajectories, we have a relation of the form

$$y''' = \alpha + \beta y'' + \gamma y''^2.$$

The center of the osculating sphere of the trajectory is then

$$X = 0,$$

$$Y = \frac{z'''}{y''z''' - z''y'''} = \frac{b + cy''}{y''(b + cy'') - a(\alpha + \beta y'' + \gamma y''^2)},$$

$$Z = \frac{-y'''}{y''z''' - z''y'''} = \frac{-(\alpha + \beta y'' + \gamma y''^2)}{y''(b + cy'') - a(\alpha + \beta y'' + \gamma y''^2)}.$$

Here y'' enters as parameter, varying from curve to curve: eliminating it, we find the locus, lying in the plane $X=0$, to be

$$\alpha Y^2 + \beta Y(1 - aZ) + \gamma(1 - aZ)^2 + Z\{bY + c(1 - aZ)\} = 0.$$

Hence for any positional force on any surface, the ∞^1 trajectories starting from a given lineal element of the surface have osculating spheres, at the common point, whose centers lie on a (general) conic in the plane normal to the element.

This conic passes through the center of curvature of the normal section of the surface determined by the given element. If the element is in one of the principal directions of the surface, the conic touches the normal to the surface.

§§ 127–130. The General Field in Space of n-Dimensions

127. Any dynamical system with n degrees of freedom may be represented by a particle in space of n dimensions. For example, an arbitrary rigid body in ordinary space is represented by a particle in six-dimensional space, and the astronomical problem of three bodies in the most general case leads to a representative particle in space of nine dimensions.

For conservative forces, or natural families, the general discussion for any dimensionality has already been given (§ 69). We shall not attempt a complete discussion for arbitrary positional forces (corresponding to that given in Chapter I for two and three dimensions). The equations of motion for an arbitrary field are

$$\ddot{x}_1 = \varphi_1(x_1, \cdots, x_n), \qquad \cdots, \qquad \ddot{x}_n = \varphi_n(x_1, \cdots, x_n).$$

We confine ourselves to the simplest questions. If the initial position and initial direction are kept fixed, and only the initial speed v is varied, what are the properties of the ∞^1 trajectories obtained? The simplest geometric result is that r_s (the rate of variation of the radius of curvature with respect to the arc length) varies as a linear function of r. The locus of the centers of the osculating spheres is a straight line, just as in the case where n is three.

128. A general curve in n-space has at each point an osculating plane, an osculating 3-flat, and so on up to an osculating $(n-1)$-flat. It is obvious that our ∞^1 trajectories have the same osculating plane since this is determined by the given initial direction and the direction of the force. It can be shown that the osculating 3-flat is also fixed; the 4-flat varies, generating a pencil; the 5-flat varies, generating a quadratic system; and so on, with more complicated variations.

129. Consider next the connection between the various curvatures and the speed.

In the plane ($n = 2$) there is only one curvature γ_1, and this varies inversely as the square of r.

In space ($n = 3$) the first curvature γ_1 varies as above, and the second curvature or torsion γ_2 remains fixed.

If $n = 4$, we have three curvatures. The laws for γ_1 and γ_2 are as above, while

$$\gamma_3 = c_1 + c_2 r^{-2},$$

where c_1, c_2 are constants (depending of course on the given initial lineal element).

If $n = 5$, we have $\gamma_1 = ar^{-2}$, $\gamma_2 = b$ (these forms are valid for any dimensions) and

$$\gamma_3 = \sqrt{c_1 + c_2 r^{-2} + c_3 r^{-4}}, \qquad \gamma_4 = \frac{d_1 + d_2 r^2 + d_3 r^4}{d_4 + d_5 r^2 + d_6 r^4}.$$

If $n = 6$, γ_3 remains the same, the numerator in γ_4 is replaced by the square root of a polynomial involving r^6, and γ_5 is given by a rational formula.

It is easy to write down the general formulas for the $n-1$ curvatures in n space. All except the first, second, and the last are irrational. These results are to be regarded as generalizations of the elementary fact (included in the formula for centrifugal force v^2/r), that the ordinary curvature varies as r^{-2}.

130. By eliminating v from any two of the formulas, we can obtain purely geometric results. For example, in space of four dimensions, $\gamma_3 = A + B\gamma_1$, where A and B depend only on the

common initial element. But in higher spaces

$$\gamma_3 = \sqrt{A + B\gamma_1 + C\gamma_1^2}.$$

This is the form required in particular in the application to the problem of three bodies, since the representative space has nine dimensions.

§§ 131–132. Interacting Particles in the Plane and in Space

131. We consider the motion of $n + 1$ particles, denoted by $M, M_1, \cdots, M_n$, moving in the plane under the action of any forces depending on the position of the particles. The differential equations of motion are then of the form

$$\begin{aligned}
\ddot{x} &= \varphi(x, y, x_1, y_1, \cdots, x_n, y_n), \\
\ddot{y} &= \psi(x, y, x_1, y_1, \cdots, x_n, y_n), \\
\ddot{x}_1 &= \varphi_1(x, y, x_1, y_1, \cdots, x_n, y_n), \\
\ddot{y}_1 &= \psi_1(x, y, x_1, y_1, \cdots, x_n, y_n),
\end{aligned}$$

and so on, where the masses—which cannot be assumed to be unity as in the case of a single particle—are absorbed with the forces in the right hand terms. From these equations the following properties may be deduced.

(1) Given the phases of $M_1, \cdots$, and the position and the direction of M, a set of ∞^1 trajectories of M is determined (one for each value of the speed). The foci of the osculating parabolas lie on a special quartic curve whose inverse with respect to the given point is a parabola tangent to the given initial line (the point of contact, however, is usually not the given point).

(2) If the speed of one of the remaining particles, say M_1, is varied, all the other initial conditions being unaltered, the parabolic locus just obtained varies. Its point of contact with the initial line remains fixed and all the ∞^1 parabolas, one for each value of the speed, are homothetic with respect to the point of tangency.

(3) The normal constructed at the common point of tangency cuts the parabola again at a distance d which varies in such a way that the square root of d can be expressed as a linear combination of the square roots of the radii of curvature of the corresponding trajectories described by the particles $M_1, \cdots, M_n$.

(4) If we preserve the phases of the particles $M_1, \cdots, M_n$, then, for each initial direction y' of M, we obtain, by (1), a certain parabolic locus. Consider the relation between the axis of this parabola and the initial direction. It is found that the initial direction y' always bisects the angle between the direction of the force acting at the given point and the direction of the axis of the parabola.

(5) Furthermore, the point where the parabola touches the initial line describes, when y' varies, a quartic curve whose inverse with respect to the given point is a conic passing through that point in the direction of the force.

It is to be observed that the statement (3) about the variation of d simplifies considerably in the case of *two* particles (that is, $n = 1$). In that case d varies directly as the radius of curvature of the trajectory described by M_1.

132. A few corresponding results for the case of any number of particles moving in space are as follows: If the speed of M is the sole arbitrary parameter, the ∞^1 trajectories of M have the same osculating plane; the torsion varies according to a linear integral function of the square root of the curvature; the locus of the centers of the osculating spheres is a cubic curve of special type.

If we assign the phases of all the particles except M_1 and assign the position and direction of M_1, then the speed of M_1, or, in consequence, the curvature of the trajectory described by M_1, is the only arbitrary parameter. There will then be ∞^1 corresponding trajectories described by M. These will of course start from the same point in the same direction with a common osculating plane and a common curvature, that is, they all have contact of the second order. The torsion varies and so does the

center of the osculating sphere. The simultaneous variation is controlled by the law that the distance from the center of the osculating sphere to the fixed center of curvature varies as a linear integral function of the radius of torsion. An equivalent statement is that the rate of variation of the radius of curvature per unit of the arc is expressed by a linear integral function of the torsion.

All these results apply in particular to the three-body problem. The present application is more concrete than that indicated in § 130, since no higher space is here introduced.*

§§ 133–141. Forces Depending on the Time. Trajectories and Space-Time Curves

133. Hitherto the force has been assumed to be independent of the time; now we consider the generalization where the force depends in any way upon the time as well as the position. Take the case of a particle moving in the plane; the equations of motion are then of the form

$$(1) \qquad \ddot{x} = \varphi(x, y, t), \qquad \ddot{y} = \psi(x, y, t).$$

From these, by differentiation and elimination, we may derive

$$(2) \qquad y''' = Py'' + Qy''^2 + Ry''^3,$$

where the coefficients are functions of x, y, y', t, namely,

$$P = \frac{\psi_x + y'\psi_y - y'(\varphi_x + y'\varphi_y)}{\psi - y'\varphi},$$

$$Q = \frac{-3\varphi}{\psi - y'\varphi}, \qquad R = \frac{\psi_t - y'\varphi_t}{(\psi - y'\varphi)^2}.$$

If we are given the initial time, position and direction, that is, the initial values of t, x, y, y', there will be a certain set of ∞^1

* Since the forces in the three-body problem are conservative, we may decompose the motions into natural families, and interpret each family in a flat space of eight dimensions. The circles of curvature at a given point will meet again; eight of them will be hyperosculating, and these will be mutually orthogonal. Cf. § 70.

trajectories, one for each value of the initial speed. The following properties are obtained:

(1) We find that the focal locus (that is, the locus of the foci of the ∞^1 osculating parabolas) is a quartic curve whose inverse with respect to the given point is a parabola which is tangent to the given direction line (the point of contact is not usually at the given point).

(2) As y' varies (x, y, t being held fixed) this point of contact describes a cubic curve whose inverse is a conic passing through the given point in the direction of the force.

(3) The initial direction of y' bisects the angle between the direction of the force and the direction of the axis of the parabola described in (1).

134. The total system of trajectories, for all initial conditions, consists of ∞^4 curves. Only in the case where the force does not depend upon the time does the system consist of ∞^3 trajectories. In the properties stated above, the initial time is kept fixed. In a certain sense then the results are not purely geometric: they would not appear in a photograph of the complete system of trajectories. This system will be represented by a certain differential equation of the fourth order; but it is not possible to carry out the requisite eliminations in explicit form, and hence the derivation of purely geometric properties involves essentially new difficulties. A complete characterization is however obtained, by projection from space curves, in §§ 136, 140.

135. There is an interesting special case in which the elimination can be carried out: namely, the problem of the motion of a *particle of variable mass* in a positional field of force. The time then appears only through the mass, so the equations of motion are of the form

$$(3) \qquad f(t)\ddot{x} = \varphi(x, y), \qquad f(t)\ddot{y} = \psi(x, y).$$

As the result of the elimination is complicated, we shall here consider only the case where the function $f(t)$, representing the mass, is of one of the special types t^4, t^2, e^t, $(\log t)^2$. The equa-

tion of the fourth order representing the trajectories is then found to be of the form

$$(4) \qquad y^{\mathrm{IV}} = Ay'''^2 + By''' + C,$$

where A, B, C involve only x, y, y', y''.

We see that the fourth derivative is a quadratic function of the third derivative. This category of equations of the fourth order arises in a number of different connections, in particular in the inverse problem of the calculus of variations, as stated in § 116. The characteristic geometric property may in the present case be stated as follows:

If the particle, whose mass varies according to one of the four laws stated, is projected into a field of force from a fixed initial position in a fixed direction at different times, with the initial speed for each time so adjusted as to cause the initial curvature of the trajectory to have a fixed value, and if for each of the ∞^1 trajectories thus obtained we construct the osculating conic (having five-point contact), the locus of the centers of these conics is a conic passing through the given conic in the given direction.

Of course not every system of ∞^4 curves having this property can be regarded as a trajectory system corresponding to equations of motion of the form considered. We do not, however, attempt a complete characterization.

136. *Space-time Curves.*—When we integrate the equations of motion, either in the special case where the forces depend only on the position

$$(1') \qquad \ddot{x} = \varphi(x, y), \qquad \ddot{y} = \psi(x, y),$$

or in the general case where the force depends also on the time

$$(1) \qquad \ddot{x} = \varphi(x, y, t), \qquad \ddot{y} = \psi(x, y, t),$$

we obtain x and y expressed as functions of t and four constants of integration. If we represent t by an ordinate perpendicular to the xy-plane, thus considering x, y, t as rectangular coordinates

in space, we obtain a certain system of ∞^4 curves in that space which we designate as *space-time curves*.*

If we project these curves orthogonally on the xy-plane, we obtain the trajectories. In the general case (1) there will be ∞^4 of these trajectories; but in the special case where the force is positional, only ∞^3 trajectories arise, since the system of space-time curves, whose number is still ∞^4, now admits the group of translations along the t-axis.

If we project the space-time curves orthogonally on the xt-plane and on the yt-plane, we obtain in each case a system of ∞^4 plane curves.

What are the properties of the system of ∞^4 space-time curves? The following two properties are characteristic:

(1). The osculating planes of the ∞^2 space-time curves through a given point go through a fixed line parallel to the xy-plane. (This line is parallel to the direction of the force acting at the projected point in the xy-plane.)

(2). If the ∞^2 space-time curves through the given point are orthogonally projected on any plane perpendicular to the xy-plane, the ∞^2 plane curves obtained are such that those which have the same tangent also have the same curvature.

Another complete characterization may be given as follows:

(3). If the ∞^2 space-time curves through a given point are orthogonally projected on either the xt-plane or the yt-plane, the ∞^2 plane curves obtained have their centers of curvature located on a special cubic of the form $t^3 = a(x^2 + t^2)$ or $t^3 = b(y^2 + t^2)$. A corresponding cubic locus will then necessarily arise by projection on any plane perpendicular to the xy-plane.

* It may be remarked that if, in problem (1), the force is multiplied by a constant c (or, what is equivalent, the mass of the particle is multiplied by $1/c$), a distinct system of ∞^4 space-time curves will be obtained. The totality of ∞^5 space curves, thus related to the ∞^1 plane problems

$$\ddot{x} = c\phi(x, y, t), \qquad \ddot{y} = c\psi(x, y, t),$$

may be generated as trajectories in a three-dimensional positional field of force. The ∞^5 curves have the four characteristic properties of a space system (§ 11) and the further peculiarity that the direction of the force is parallel to the xy-plane.

137. Consider the ∞^4 curves in say the xt-plane. These are the curves representing graphically the relation between the abscissa x and the time t. By eliminating y from the set (1), we obtain a relation of the form

$$x^{\text{IV}} = A\dddot{x}^2 + B\dddot{x} + C,$$

where A, B, C involve only x, $\dot{x}$, $\ddot{x}$ and the independent variable t. The fourth derivative is thus always quadratic with respect to the third derivative. Hence, by § 116, we have this result:

In the xt-plane (or, more generally, in any plane perpendicular to the plane xy in which the motion actually takes place), the ∞^1 curves having any element of curvature in common are such that the locus of the centers C'' of their osculating conics (constructed at the common point) is a conic passing through the common point in the direction of the common tangent.

As indicated above, the ∞^4 curves in the xy-plane, that is, the trajectories, do not usually enjoy this simple property. Even in the case where the time enters only through the mass, the locus of the centers of the osculating conics may be of any degree of complication. Its shape depends on the law of variation of the mass. Only for the special laws stated at the bottom of page 112, together with certain combinations of them, is the equation of the trajectories of the quadratic type.

138. It is possible to obtain additional general properties of the xt-system, describing how the locus conic, corresponding to a curvature element, changes when the element changes. For the coefficients A, B, C determining the position of the conic have the following forms: A does not involve $\dot{x}$, B is linear and integral in $\dot{x}$, C is quadratic and integral in $\dot{x}$. Hence these results:

If the curvature element is varied, at the given point O, in such a way that the second derivative $\ddot{x}$ is constant, so that only $\dot{x}$ varies, the center C'' of the corresponding locus conic describes a new conic.

At the same time a certain two-to-one correspondence arises between the initial direction of the element and the direction of the line joining O to the center C''.

139. A clearer picture is perhaps obtained by changing the notation to correspond with the usual x, y, z notation for rectangular coordinates in space. It is then desirable to lay off the time on the x-axis, since this is the independent variable. The actual motion then takes place in the yz-plane, and the differential equations of motion are

$$\frac{d^2y}{dx^2} = \varphi(x, y, z), \qquad \frac{d^2z}{dx^2} = \psi(x, y, z).$$

The curves in space x, y, z are then the space-time curves. Their projections on the yz-plane are the trajectories (whose explicit properties have not been derived). Their projections on the xy-plane (or on the xz-plane, or on any plane parallel to the z-axis) are curves whose properties have just been stated (§§ 137, 138). The differential equation in the xy-plane is

$$y^{\text{IV}} = \alpha y'''^2 + (\beta_1 + \beta_2 y')y''' + (\gamma_1 + \gamma_2 y' + \gamma_3 y''^2),$$

where the coefficients involve only x, y, and y'.

140. We have not attempted a complete direct characterization of the systems of curves arising in any one of the coordinate planes. Such a characterization has however been given (§ 136) for the system of ∞^4 space-time curves. Indirectly this really solves all the problems. A system of curves in the plane can be regarded as trajectories of a force depending on time and position if and only if the curves can be obtained by orthogonal projection from some system of ∞^4 curves in space having the properties (1) and (2) of § 136. If, furthermore, the space system is invariant under translation perpendicular to the given plane, the plane system, then consisting of only ∞^3 curves, belongs to a positional field.

141. For any force depending on time and position

$$\ddot{x} = \varphi(x, y, t), \qquad \ddot{y} = \psi(x, y, t),$$

the number of space-time curves is always ∞^4. When we project

these on the xy-plane, to obtain the trajectories, the number is usually ∞^4. The number reduces to ∞^3 if the force is positional but does not vanish; in the latter case the trajectories are merely the ∞^2 straight lines.

In the xt-plane the usual number of curves is ∞^4. The only exception arises when the function φ is free from the variable y. In this case the xt-curves all satisfy the equation of second order $\ddot{x} = \varphi(x, t)$ and therefore their number is only ∞^2. Similar statements hold of course for the yt-plane.

Consider, as a single example, gravity, taken as uniform and acting in the vertical xy-plane. The equations of motion are

$$\ddot{x} = 0, \qquad \ddot{y} = g.$$

The xyt-curves are

$$x = at + b, \qquad y = \tfrac{1}{2}gt^2 + ct + d,$$

a certain family of ∞^4 parabolas in space. The xt-curves are ∞^2 straight lines. The yt-curves are ∞^2 parabolas. The xy-curves (that is, the trajectories) are ∞^3 parabolas

$$y = \alpha x^2 + \beta x^2 + \gamma.$$

It is to be observed that if the gravity constant g is changed, the new problem, while giving the same trajectories, gives a distinct family of xyt-curves. If g takes all possible values, the totality of space-time curves obtained is formed of ∞^5 parabolas (namely, those whose axes are parallel to the t-axis). These curves, in accordance with the general statement made in the footnote on page 114, are the trajectories of a positional field in space, the generating force being constant and acting in the t-direction.

All the results can be extended so as to apply to the four-dimensional space-time curves depicting motion in ordinary space.